Set Valued Dynamic Treatment Regimes

by

Tianshuang Wu

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Statistics)
in The University of Michigan
2016

Doctoral Committee:

Professor Susan A Murphy, Chair
Professor Xuming He
Assistant Professor Ambuj Tewari
Associate Professor Lu Wang

To my parents Jing and Wei

To my wife Yingyi and my daughter Audrey

ACKNOWLEDGEMENTS

I would like to give my thanks to my advisor Prof. Susan Murphy. For the past five years, she turned me from a student whose only advantage was taking exam to a researcher by teaching me how to discover a scientific problem, evaluate it and find ways to deal with it. However, the help and inspiration from her includes far more than knowledge and research. As a role model, she showed me almost all the virtue of a researcher: solid mathematic background, the enthusiasm of diving into the problem, the ability of communicating with people both inside and outside Statistics field and work with people in the research lab. I learned these from her not from her word, but her act. I feel really lucky to have her as my mentor. Everything that I learned from her has been, and will always stay with me and help me for the rest of my life.

I would thank my committee members, Prof. He, Prof. Tewari and Prof. Wang, for their kind help and advise during these years. I could always gain fresh and new ideas when discussing with them. Also I would thank Prof. Almirall and the former postdoc Ashkan Ertefaie, for their cooperation for the paper we wrote together. They are both great researchers and good friends and I have had a good time working with them. Of course, I would thank everyone in our lab. They are perfect friends, colleagues and I'm grateful that I have them around during my time at University of Michigan.

Finally I would give my most sincere thanks to my parents Wei and Jing, my wife Yingyi and my lovely daughter Audrey. None of my work could have been

done without their support and encouragement, and none of my work would be so meaningful to me without them.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF ABBREVIATIONS

FWER Family-Wise Error Rate

LRT Likelihood ratio test

MCB Multiple Comparisons with the Best

MCC Multiple Comparison with the Control

SMART Sequential Multi-Assignments Randomized Trials

UIP Union-intersection principle

ABSTRACT

Set Valued Dynamic Treatment Regimes

by

Tianshuang Wu

Chair: Susan Murphy

Dynamic Treatment Regimes (DTR)s are composed of sequences of decision rules, one per stage of treatment. Each decision rule inputs patient information and outputs a single recommended treatment. While the majority of present studies are focused on finding the optimal DTR, we take another approach. Instead of trying to determine the true best DTR, we aim to construct a set of DTRs such that the true best DTR is contained in this set with a desired probability. The reasons are as follows: (1) Usually we do not have enough data to identify the best DTR and (2) we want to give patients and clinicians more options. To discuss the second reason in more detail, patients and clinicians might have treatment preferences related to cost, side effects or convenience, etc. Thus, our goal is to provide a recommended set of DTRs, such that the DTRs contained in the set are those we cannot distinguish from the best, while the DTRs we exclude are those that are certain to be inferior with high confidence. This idea comes from decision support: we do not tell patients and clinicians what to do; we do not offer treatments known to be inferior. Rather we offer a set of treatments that excludes inferior treatments. In this thesis we develop a set valued DTR in which the decision rules at each stage can output a set of treatments. Second we develop

an approach for constructing a recommended set of DTRs. In the appendix we prove the relevant theorems.

CHAPTER I

Introduction

Comparisons of treatments is always crucial when we have more than one available candidate at hand at a certain situation, e.g., at a certain time point for a certain kind of patients. For the comparison of two populations, there are a lot of well developed methods from parametric methods like t-test to non-parametric methods like U-test. However, in practice, we often face the situation when more than two treatments are available at hand and we would like to compare them together. One naive way is to compare them pairwise and produce $\binom{N}{2}$ confidence intervals of the difference s of the effects between all pair of treatments. Naturally this approach fails to consider the Family-Wise Error Rate (FWER) (see *Shaffer* (1995) for more details). Rescues have been made by different approaches, like the Bonferroni correction proposed by *Bonferroni* (1936) and developed by *Dunn* (1959, 1961),the Sidak correction which is credited to *Šidák* (1967) by *Seidler et al.* (2000). If there is one "standard effect" that serves as a control and we want to compare the effect of each treatment with this control, a simultaneous comparison method called Multiple Comparison with the Control (MCC) (*Dunnett*, 1955) can be applied. If we are only interested in whether all effects of the treatments are equal, ANOVA (*Box et al.*, 1954a,b) could be applied. But if we want further know the relationship between the effects of the treatments, the Multiple Comparisons with the Best (MCB) method proposed by *Gupta* (1965) and

developed by *Hsu* (1996) can be applied. This method is the main technique of many procedures in this thesis and will be introduced in more details in the corresponding sections.

Besides the effect of treatments, sometimes we are also interested in comparing two or more sequences of decision rules that provide sequences of treatments over time. A sequence of decision rules that adapts over time, which take patients characteristics at each time point as input and output a recommended treatment, have different names in different fields, like Dynamic Treatment Regimes (DTRs) (*Robins*, 1986, 1989, 1993, 1997, 2004), adaptive treatment strategies (*Lavori and Dawson*, 2000, 2008) (*Murphy*, 2005a) (*Thall et al.*, 2000, 2011), treatment policies (*Lunceford et al.*, 2002) (*Wahed and Tsiatis*, 2004, 2006) or adaptive interventions (*Almirall et al.*, 2014). *Lei et al.* (2012) provided a detailed introduction of DTRs. There are a lot of scientific questions that can be asked, like finding the best DTR among several given DTRs, the comparison of DTRs with the same or different initial treatments etc. *Murphy* (2005a) proposed a clinical design named Sequential Multi-Assignments Randomized Trials (SMART) that can obtain data and answer these questions efficiently. To analyze the data from a SMART, Q-learning technique (*Watkins and Dayan*, 1992; *Murphy*, 2005b) can be applied. There are a lot of variation of this technique, examples can be found in (*Murphy*, 2003; *Robins*, 2004; *Blatt et al.*, 2004; *Moodie et al.*, 2007; *Henderson et al.*, 2010; *Zhao et al.*, 2009). However, they all suffer from the same issue called non-regularity, which will be discussed in detail in chapter III. A number of papers have discussed this problem including (*Robins*, 2004; *Moodie et al.*, 2007; *Henderson et al.*, 2010; *Chakraborty et al.*, 2009; *Moodie and Richardson*, 2010; *Laber et al.*, 2010).

In practice, we will often encounter different scenarios that we need to make recommendation among several treatments at a time point, or among several DTRs at the beginning of a session. Traditionally, researchers used data from clinical trials to

estimate the "effects" of each treatment or DTR, and recommend patients or clinicians with the one with the highest estimated effect. We are trying to provide another approach. As those treatments or DTRs that do not have the highest estimated effects, hence abandoned, might have other advantages like lower burden of side effects or financially, or better agreement with patients schedule. So we do not want to given them up that easily since their having lower estimated effect than the chosen one might be merely by chance. So following the idea of *Horrace et al.* (2000), we would like to construct a recommended *set* of treatments of DTRs, such that those contained in the set are those statistically we cannot distinguish from being the *best*, while those we excluded from the set are those we are statistically certain to be inferior than the *best*. Our work follows the idea of clinical decision support system in the sense that we do not tell clinicians or patients what to do, because usually we do not have enough statistical evidence to do so, but rather tell them what not to do.

The rest of this thesis are organized as follows. In chapter II we will focus on the situation where there are several one-stage treatments and we would like to construct a recommended set. In that chapter we will first review the existing methods and then introduce two new methods that are more "efficient" than the popular MCB method. In chapter III we will discuss the construction of treatments when we have more than one stages. The key part is the new definition of the effect of a non-final stage treatment and the technique of constructing the confidence interval of the differences of the effects. Chapter IV will deal with the situation where we would like to construct a set of DTRs. This is the joint work with some other researchers and the result has been published in Biostatistics. The final discussion and future work will be in chapter V. The proof as well as the graphs can be found in the Appendix.

CHAPTER II

Comparison methods

2.1 Introduction

As has been mentioned in chapter I, when there are more than one treatment available, instead of trying to find the treatment with the best "effect", we would like to construct a set of treatments such that the best one is contained with high probability. But if this is our only goal, then naively including all the treatments in a set will work. Thus we are seeking other properties of this set.

We would like to start with a simple setting, i,e., the estimators of the treatments are independent and have the same variance. Let's consider N populations indexed by $\{1, 2, \cdots, N\}$; corresponding to these N treatments we have observations $\{X_{ij}\}$ where $i = 1, \cdots, N$ and $j = 1, \cdots, n$ and n is the common sample size of each population. We assume that observations are independent to each other and observations from the ith population follow $N(\theta_i, \sigma^2)$, where θ_i is the mean of the ith population . Either the common variance σ^2 is known or can be consistently estimated by $\hat{\sigma}^2$ (usually the pooled sample variance). Our goal is to find a set $\hat{S}$, based on the data, that contains the best (i.e. with largest mean) index with probability more than a given level $1 - \alpha$. Let $\theta_{(1)} \leq, \cdots, \leq \theta_{(N)}$ be an ordered sequence, then we need $P((N) \in \hat{S} | \theta_1, \cdots, \theta_N) \geq 1 - \alpha \quad \forall \theta_i' s$. Let $\hat{\theta}_i = \frac{1}{n}\Sigma_{j=1}^{n} X_{ij}$ $(i = 1, ..., N)$ be the sample average of the ith indexed population and thus a natural estimator of θ_i,

and $\hat{\theta} = (\hat{\theta}_1, ..., \hat{\theta}_N)^T$. Let $\delta_i = (\theta_i - \theta_1, \theta_i - \theta_2, \cdots, \theta_i - \theta_N)$ be an N-1 dimensional vector denoting the differences between the ith mean and other means, let $\hat{\delta}_i$ be the corresponding estimator with θs replaced by $\hat{\theta}$s.

Here, we are comparing several populations instead of two. One idea is to compare each pair of them, and screen out those population who are 'significantly worse' than some other populations. This means that we are constructing $\binom{N}{2}$ confidence intervals in total, each for a pair of true difference, and for each index i, it is excluded from the set if and only if all the $N-1$ confidence intervals that involve θ_i suggest that θ_i is lower. As has been discussed by *Hsu* (1996), this approach suffers from the common shortcoming of Bonferroni method: the lack of power. i.e., the probability of excluding true inferior treatments is very low even if there is a moderate difference. The second idea is that we consider all the means as one vector. This vector θ lies in an N dimensional space which can be divided into N parts, each of which is the set where one particular θ_i is larger than or equal to all other θ_js. If we fail to reject that θ lies in the space where θ_i is largest, we keep index i in our confidence set. Many methods we introduce below follow this idea. Further more, as what we concern is only the differences between them, not the exact value. So we can consider the difference between one of the population and the others. Thus what we concern is an $N-1$ dimensional vector δ. In such a way we can reduce the dimension of the vector of interest by one and obtain more power.

The rest of this chapter is organized as follows, section 2.2 will contain a review of several existing comparison methods and their pros and cons. In section 2.3 we will introduce two new methods as supplements of Multiple Comparisons with the Best technique.

2.2 Review of the comparison methods

The first two of the methods reviewed in this section use the technique of Hypothesis testing, which can be inverted to the problem of constructing confidence sets. We let index i into $\hat{S}$ if and only if we fail to reject the null hypothesis

$$H_0 : \qquad \theta_i \geq \theta_j \quad \forall j \in \{1, 2, \cdots, N\} \tag{2.1}$$

The null space is that the difference vector δ lies in the place where the ith element of θ is the largest. And the alternative space is the complement of H_0

2.2.1 Union-intersection principle (UIP) method

Union-intersection principle (UIP) is introduced by Roy *Roy* (1953), the idea is that we decompose the alternative as the union of $\{H_i\}_{i=1}^K$, and for each H_0 v.s H_i we have a test, then we accept H_0 iff we accept all the null of subtests. In other words, we reject H_0 iff we accept some H_i. The rationale of this decomposition is that although the null space is convex, the alternative space is not. Thus if we directly use likelihood ratio test (as we would introduce later), there will be problems. We will talk about this in detail in that part. Now we decompose H_1 into several parts and it is easier to compare each of them against H_0.

What Gupta uses here is a variation of UIP, which is mentioned by *Sen* (2007). Here we write

$$H_0 = \bigcap_{j \in \mathcal{F}} H_{0j}, \quad H_1 = \bigcup_{j \in \mathcal{F}} H_{1j} \tag{2.2}$$

where $\mathcal{F}$ is a suitable index set, and for each $j \in \mathcal{F}$, there exists a suitable test for testing H_{0j} against H_{1j}. And again we accept H_0 iff we accept all the H_{0j}.

Gupta's UIP method is introduced by *Gupta* (1956). Here $\mathcal{F} = \{1, 2, \cdots, N\}$

and $H_{0j} = \{\theta_i \geq \theta_j\}$,$H_{1j} = \{\theta_i < \theta_j\}$. For each j the test statistic is $\hat{\eta}_{ij} = \sqrt{n}(\hat{\theta}_i - \hat{\theta}_j)/\sigma_{ij}$ where σ_{ij} is the standard deviation of $X_i - X_j$. If we don't know it, we replace it by the pooled sample standard error $\hat{\sigma}_{ij}$. So the test statistic for population i is $\hat{\eta}_i = (\sqrt{n}(\hat{\theta}_i - \hat{\theta}_1)/\sigma_{i1}, \sqrt{n}(\hat{\theta}_i - \hat{\theta}_2)/\sigma_{i2}, \cdots, \sqrt{n}(\hat{\theta}_i - \hat{\theta}_N)/\sigma_{iN})$. We accept H_{0j} if $\hat{\eta}_{ij} \geq -d_{ij}$, where $\{d_{ij}\}_{j=1,j\neq i}^{N}$ is determined to guarantee the $1-\alpha$ size, i.e. for i fixed,

$$P(\hat{\eta}_{ij} \geq -d_{ij}\ \forall j\ |\theta_i = \theta_j\ \forall j) = 1 - \alpha \tag{2.3}$$

As this is the worst case scenario. To explain, our ultimate goal is that

$$P(\hat{\eta}_{ij} \geq -d_{ij}\ \forall j\ |\theta_i \geq \theta_j\ \forall j) \geq 1 - \alpha \tag{2.4}$$

As enlarging θ_j (for $j \neq i$) will make $\hat{\eta}_{ij}$ smaller, so the worse case, i.e. the smallest value of the left hand side of (2.4) for fixed $\{d_{ij}\}$ happens when $\theta_i = \theta_j\ \forall j$. i.e.

$$P(\hat{\eta}_{ij} \geq -d_{ij}\ \forall j\ |\theta_i = \theta_j\ \forall j) \geq 1 - \alpha \tag{2.5}$$

But we can always enlarge the power (i.e. make all $\{d_{ij}\}$ smaller, say, by timing a number γ a little smaller than 1 to all $\{d_{ij}\}$, thus the left hand side of (2.5) will be smaller but as normal distribution is continuous so by carefully choosing γ we can let it be still no smaller than $1-\alpha$) while still preserving size α. So the optimal set of $\{d_{ij}\}$ satisfies (2.3), in the sense that we cannot shrink any d_{ij} without enlarging at least one of the other d_{ij}.

And thus the rule of forming $\hat{S}$ is that

$$\hat{S} = \{\ i \mid \hat{\eta}_{ij} \geq -d_{ij}\ \forall j \neq i\} \tag{2.6}$$

Under the case that all means are the same, $\hat{\eta}_i$ (we drop the term $\hat{\eta}_{ii}$ to make $\hat{\eta}_i$ N-1 dimension since $\hat{\eta}_{ii}$ is always zero) is distributed as a multivariate normal (if we know

the variances) or t (if we don't know the variances) distribution with no unknown parameter. Denote it as $Z = (Z_1, Z_2, \cdots, Z_N)$ (Note that this is N-1 dimension as we don't have the term Z_i) so (2.3) becomes

$$P(Z_j \geq -d_{ij},\ \forall j \neq i) = 1 - \alpha \tag{2.7}$$

In principle any $\{d_{ij}\}$ satisfying (2.7) can form an $\hat{S}$ with desired coverage probability. Gupta chose the case that $d_{ij} \equiv d_i$.

Edwards and Hsu (1983) showed that following (2.6), not only we have $1 - \alpha$ probability of covering $(N) := \arg\max_{i=1,\cdots,N} \theta_i$, but also have an estimate of $\theta_{(N)} - \theta_i$ for each i.i.e

$$P((N) \in \hat{S}, \theta_{(N)} - \theta_i \leq U_i) \geq 1 - \alpha \tag{2.8}$$

for some U_i determined by the data. And this can be done with the following knowledge of variance matrix, either we completely know it or at least we know the structure of it, i.e. we know that variance matrix $\Sigma = \sigma^2 C$ where C is a known matrix and σ is a (possibly unknown) scalar (This is called general MCB, comparing with standard MCB, which is that $\hat{\theta}_i$s are independent. Here MCB stands for multiple comparisons with the best).

We can see this in another point of view, we are constructing a 'one-sided confidence interval' of the true δ base on (2.6), and recall we are dividing R^{N-1} into N subspaces such that each of them represents the area where a certain θ_i is the largest. From (2.6) it's equivalent that we let index i into the confidence set if and only if the 'confidence interval' has non-empty interception with the area where θ_i is the largest. We will see another kind of 'confidence interval' in the section of likelihood ratio test.

If we estimate treatment effects using only the average of effects from samples, then usually the estimators are independent, but if they are combined with some other covariates with unknown coefficients, then we need to estimate both the treatment

effects and those unknown coefficients, thus leads to dependency. Below are detailed explanation. This example comes from *Horrace et al.* (2000).

We consider panel data regression model

$$y_{it} = \theta_i + x'_{it}\beta + \epsilon_{it} \quad i = 1, \cdots, N,\ t = 1, \cdots, n \tag{2.9}$$

Here y_{it} is the tth observation of the ith population, θ_i is the effect of the ith treatment, x_{it} is some feature covariates of the tth person given the ith treatment and ϵ_{it} is the error. For simplicity, we assume that for each population there are equally n observations and the error ϵ_{it} are iid normal with mean 0 and variance σ^2.

If β were known, then we could write $(y_{it} - x'_{it}\beta) = \theta_i + \epsilon_i$, then this reduces to standard MCB. If β is not known but we have a consistent estimator $\hat{\beta}$ from the regression in deviations from individual means (i.e. by regressing $y_{it} - \bar{y}_i$ on $x_{it} - \bar{x}_i$) and then define $\hat{\theta}_i = \bar{y}_i - \bar{x}_i\hat{\beta}$. This leads to the following expression for $\hat{\theta}_i$:

$$\hat{\theta}_i = \theta_i + \bar{\epsilon}_i - \bar{x}_i(\hat{\beta} - \beta) \tag{2.10}$$

$\hat{\theta}_i$ can be used as an estimator for θ_i, it's unbiased and consistent if $\hat{\beta}$ is unbiased and consistent for β. Let $\tilde{x}$ be the $Nn \times K$ matrix (K is the dimension of x_{it}) of x's expressed in deviations from individual means, so that its typical row is of the form $x_{it} - \bar{x}_i$. Let $\bar{x}$ be the $N \times K$ matrix whose ith row is $\bar{x}_i$. Then the variance matrix of $\hat{\beta}$ is $\sigma^2(\tilde{x}'\tilde{x})^{-1}$ and variance matrix of $\hat{\theta}$ is $\sigma^2 C$ where $C = I_N/n + \bar{x}(\tilde{x}'\tilde{x})^{-1}\bar{x}'$. This leads to general MCB.

If we consider a more general case. In model (2.9), the 'treatment effect'(θ_i) is the same for all members taking the same treatment. We can generalize this to the model that the treatment effect is also effected by characteristics of the individual.

That is we are considering the model

$$y_{ij} = x'_{ij}\alpha + x'_{ij}\beta_i + \epsilon_{ij} \quad i = 1, \cdots, N,\ j = 1, \cdots, n \tag{2.11}$$

So here α is a fixed vector (with length K, suppose) which can be considered as 'fix effect' for person with characteristic x_{ij}. β_i (with length K) is the characteristic of treatment i. So $x'_{ij}\beta_i$ will be the treatment effect for individual with x_{ij} and taking the ith treatment.

Now let's define $A_k(i,j)$ (k=1,...N) as the indicator, it's 1 if $i = k$ and 0 otherwise. It means whether the ijth person has taken treatment k.

Let $\tilde{x}_{ij} = (x'_{ij}, A_1(i,j)x'_{ij}, \cdots, A_N(i,j)x'_{ij})$ be a 1 by $(N+1)K$ vector and $X = (x'_{11}, \cdots, x'_{Nn})'$ be an Nn by $(N+1)K$ design matrix . $Y' = (y_{11}, \cdots, y_{Nn})$, $\beta = (\alpha', \beta'_1, \cdots, \beta'_N)'$ and $\epsilon = (\epsilon_{11}, \cdots, \epsilon_{Nn})'$ be outcome, total coefficient and error vectors, respectively. So we have the general linear regression model $Y = X\beta + \epsilon$.

So the OLS solution is $\hat{\beta} = (X^TX)^{-1}X^TY$. For a (possibly new) person with characteristic x_0, the treatment effect for him from treatment i is $x'_0\beta_i$ and the corresponding estimate is $\hat{\theta}_i = x'_0\hat{\beta}_i$. So the vector we are interested in is $\hat{\theta} = (\hat{\theta}_1, \cdots, \hat{\theta}_N) = (x'_0\hat{\beta}_1, \cdots, x'_0\hat{\beta}_N) = X_0\hat{\beta}$. Where X_0 is an N by $(N+1)K$ matrix, its ith row is constructed as follows, first a zero vector with length iK, follows by x_0 and then another $(N-i)K$ zeros. Thus we have $(x'_0\hat{\beta}_1, \cdots, x'_0\hat{\beta}_N) = X_0\hat{\beta}$. (Recall $\hat{\beta} = (\hat{\alpha}', \hat{\beta}'_1, \cdots, \hat{\beta}'_N)'$).

Assuming $\epsilon_{ij} \sim iid\ N(0, \sigma^2)$. We have $\hat{\theta} = X_0\hat{\beta} = X_0(X^TX)^{-1}X^TY$, $Var(\hat{\beta}) = \sigma^2(X^TX)^{-1}$ and $Var(\hat{\theta}) = Var(X_0\hat{\beta}) = \sigma^2X_0(X^TX)^{-1}X_0^T$. Now we finish the construction of generalized MCB.

Gupta's method is a special case of *Seal* (1955). To decide whether a population is selected into the confidence set, instead of just from the difference between the target sample average and the largest sample average (Gupta's setting), Seal considered

the difference between the target sample average with the weighted average of all other sample averages. Below is the detailed construction: in his paper, he first assumed that the sample size (denoted by n as usual, same for all populations) is so large that we can treat the estimated variance as the true variance (he assumed that all populations have the same variance and independent). Denote s as the common standard deviation of N normal distributions, and x_i be the sample average of the ith population, and $x_{(1)} < \cdots < x_{(N)}$ be the ranked observation (as normal distribution is continuous we assume no equal observations).

To find the 'critical number', let y_i $(i = 0, 1, \cdots, N-1)$ be N observation from $N(0, s^2)$ and let $y_{(1)} < \cdots < y_{(N-1)}$ be $N-1$ ranked observations among $y_1, \cdots, y_{N-1}$. The $y_{(i)}$s will define another set of random variables $Y_{(i)}$ $(i = 1, \cdots, N-1)$. Let $t_\alpha(c_1, \cdots, c_{N-1})$ denote the upper $100\alpha\%$ point in the probability density function of

$$t(c_1, \cdots, c_{N-1}) = \frac{\sum_{i=1}^{N-1} c_i Y_{(i)} - Y_0}{s} \tag{2.12}$$

The class $\mathcal{C}$ of decision rules $D(c_1, \cdots, c_{N-1})$ with $c_i \geq 0, i = 1, \cdots, N-1, \sum_{i=1}^{N-1} c_i = 1$ is defined as follows: reject any observation x_0 from the given observations $x_i, i = 0, 1, \cdots, N-1$ if

$$\sum_{i=1}^{N-1} c_i x_{(i)} - x_0 > s t_\alpha(c_1, \cdots, c_{N-1}) \tag{2.13}$$

and accept otherwise. Note that if we let $c_{N-1} = 1$ and all other c_i's equal to 0, then this is Gupta's method. In his paper, Seal discussed another special case when all c_is are equal, i.e. $c_1 =, \cdots, c_{N-1} = 1/(N-1)$, (he denoted the decision rule in this case as $\bar{D}$). First, he proved that this setting minimized the variance among all $\sum_{i=1}^{N-1} c_i Y_{(i)}$ such that $\sum_{i=1}^{N-1} c_i = 1$. In the case that all but one population have the equal mean, i.e. $\theta_1 = \theta_2 = \cdots = \theta_{N-1} = \theta_N + \delta$ where δ can be positive or negative, naturally we want to (a) maximize the probability of selecting the Nth population if $\delta > 0$ and (b) minimize the probability of selecting the Nth population if $\delta < 0$. Seal proved that

$\bar{D}$ can achieve both requirements (a) and (b) among the set of rules $D(c_1, \cdots, c_{N-1})$, the optimality means that under this rule, the the maximum probability of including the best and excluding the worst is achieved, among the set of rules $D(c_1, \cdots, c_{N-1})$. The reason why Seal considered this scenarios is that he wanted a test that contains the best and reject the worse at the same time, with high probability, while our goal is just including the best.

If we again go back to confidence interval point of view, Seal's confidence interval generally has different shape comparing to Gupta's. Gupta's interval is a cube (in θ's R^N space) or a parallelogram (in δ's R^{N-1} space) and the surface of the interval is parallel to the boundary of the areas (recall that different areas represent spaces where a certain θ_i is the largest). Generally Seal's confidence interval is a parallelogram and not parallel to the boundaries. Thus it is lack of power generally.

Comparing with Gupta's method, we can see that Seal's $\bar{D}$ method gives more chance (to be selected into the confidence set) to the true best population if it does not perform the best (i.e. has the largest sample average), but at the same time of course gives more chance to other populations as well. So generally Seal's set size will be larger than Gupta's. And Seal's method works only well in his special scenario (all means but one are equal), which is not a common case. And it will perform very bad if only one population has very low mean while the differences between others are relatively small. Because in this case each population (except the one with lowest mean) will be selected into the confidence set with high probability as its sample average will be larger than the average of all other sample averages. Keep in mind that our goal is that under the restriction of containing the best population with high probability, try to get as small a set as possible (here the measure of a set is how many elements it contains).

Based on Seal's work, *Hsu* (1985) considered another similar selection rule. In his paper he considered the situation when we have different sample sizes n_i for the ith

population. Naturally instead of $c_i = 1/(N-1)$ in Seal's case, we have

$$c_i = \frac{n_i}{\sum_{j \neq i} n_j}$$

Also he considered the case that we don't know the common variance, then the s in (2.13) is replaced by the estimated standard deviation $\hat{s}$ and the critical value t_α changes correspondingly to satisfy (2.12) with s replaced by $\hat{s}$. Similar to Seal's method, this rule maximizes the probability of including the best as well as excluding the worst.

2.2.2 Likelihood Ratio Test (LRT)

LRT is always a popular technique. *Perlman et al.* (1999) favored it to the many 'Uniformly more powerful' tests. In their paper, they listed several tests that carefully enlarged the rejection region (while preserving the $1-\alpha$ level) and hence trivially enlarged the power. But these new rejection regions violated the intuition of hypothesis testing. For example, in the problem about the mean of a normal distribution with unknown variance, $H_0 : |\mu| > 1$ and $H_a = H_0^c$. The rejection region is in figure 2.1, where x is observed sample average and s the sample standard deviation. LRT uses just the small triangular region A as the rejection region, while the new test adds the region Q as well. Thus, if we observe $\bar{x} = 0$ and $\hat{s} = 10^{100}$, then using this new test this observation provides strong evidence *against* the null. Their reason is that with such large deviation, no one can distinguish between $\mu = 1$ and $\mu = -1$, let along between H_0 and H_1.

Another example can be found in the work of *Berger* (1989). He considered a problem similar to our multiple comparisons. His H_0 is that the mean vector δ lies outside the region where all the component of it are positive. And $H_1 = H_0^c$. (In our setting, we let δ be the difference between ith mean and others', then our H_0 and

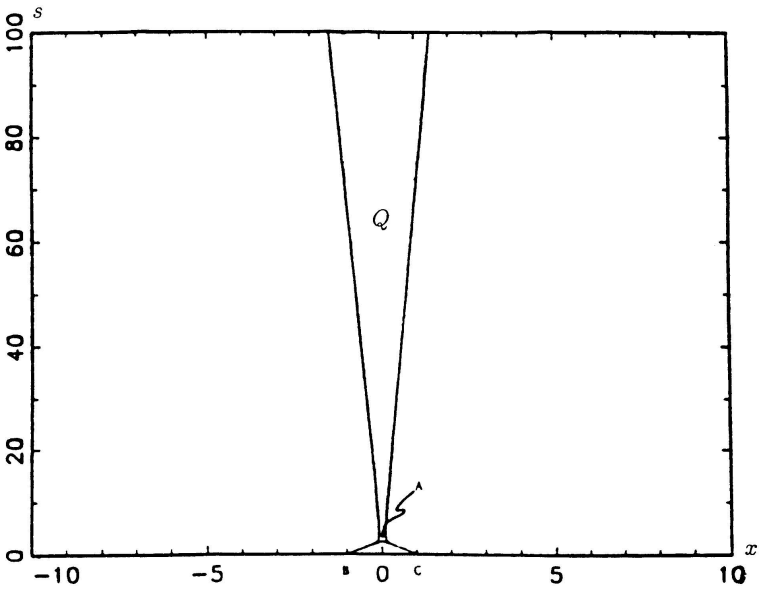

Figure 2.1: The rejection regions of Perlman's test.

H_1 are Berger's H_1 and H_0, respectively).He carefully enlarged the rejection region of LRT while still preserving the level $1-\alpha$ confidence level. Thus it trivially enlarged the power uniformly. But Perlman criticized this method as its rejection region has zero distance from the null space, i.e. we can have sample mean infinitely close to the null space but still gets rejected. In his other new test, the rejection region even contains some area of the null space. See figure 2.2, where the rejection region of Berger's type 1 test is $R_1 \bigcup R_2 \bigcup \cdots \bigcup R_5$, of type 2 test is $R_1 \bigcup R_2 \bigcup \cdots \bigcup R_9$, and of likelihood ratio test is just R_1. So it might happen that you reject the null of not all coordinates are positive when your observation averages are all negative.

Now we focus on index i, for simplicity we assume that we know the variance of the populations (and they are equal, so without loss of generality, equal to 1). The null hypothesis is the same as in (2.1) and the alternative is H_0^c. Assume we have equal samples from each population, and let $Y_j = (Y_{j1}, \cdots, Y_{jN}) = (X_{ji} - X_{j1}, \cdots, X_{ji} - X_{jN})$, $j = 1, 2, \cdots, n$ be the difference between the ith population and the others, in the jth sample, note that Y_j is N-1 dimensional. And Y_j is again multivariate normal with mean vector $\delta = (\delta_1, \cdots, \delta_N) = (\theta_i - \theta_1, \cdots, \theta_i - \theta_N)$ and the known

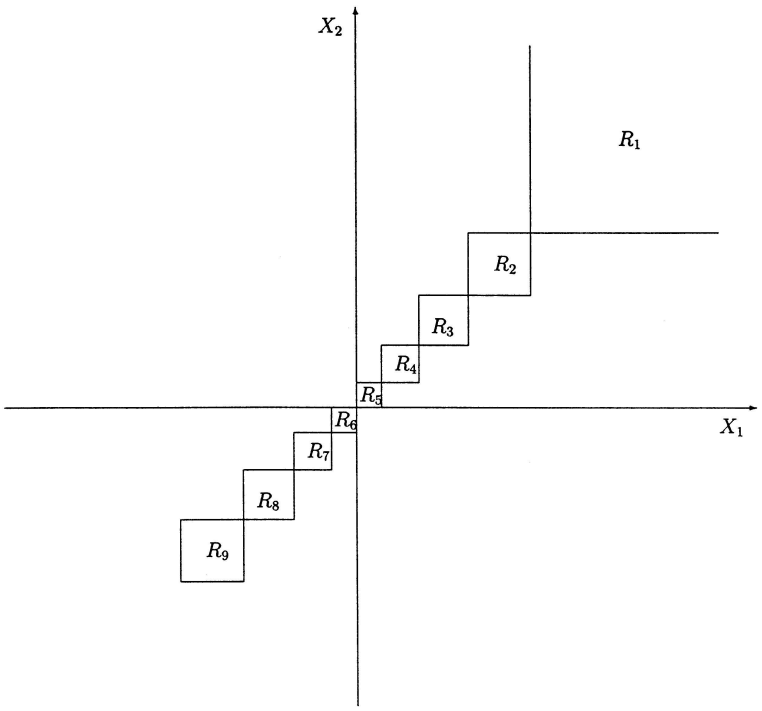

Figure 2.2: The rejection regions of Berger's two new tests.

variance matrix here denoted by Σ_j, the variance matrix of $(X_i - X_1, \cdots, X_i - X_N)$. In general, if the variance matrix of X is Σ_X then $\Sigma_j = D_j \Sigma_X D_j$ where D_j has -1 in its jth column, and the remaining columns are the columns (in order) of $-I_{N-1}$. A special case is that our Xs are independent ($\Sigma_X = I$) then Σ_j is a matrix with diagonal elements 1 and other elements 1/2. So the likelihood function is proportional to

$$\exp(-\frac{1}{2}\sum_{j=1}^{n}(Y_j - \delta)^T \Sigma^{-1}(Y_j - \delta)) \tag{2.14}$$

Under the null, the δ that maximize (2.14) is $\tilde{\delta} = (\tilde{\delta}_1, \tilde{\delta}_2, \cdots, \tilde{\delta}_N) = (\hat{\delta}_1 1(\hat{\delta}_1 \geq 0), \hat{\delta}_2 1(\hat{\delta}_2 \geq 0), \cdots, \hat{\delta}_N 1(\hat{\delta}_N \geq 0))$.

Under the whole space, the δ that maximize (2.14) is $\hat{\delta}$, where $\hat{\delta} = (\hat{\delta}_1, \cdots, \hat{\delta}_N) = (\hat{\theta}_i - \hat{\theta}_1, \cdots, \hat{\theta}_i - \hat{\theta}_N)$.

So do the division and finally we have the rule, is that we accept H_0 if

$$f(Y) = \sum_{j=1}^{n} \{(Y_j - \tilde{\delta})^T \Sigma^{-1} (Y_j - \tilde{\delta}) - (Y_j - \hat{\delta})^T \Sigma^{-1} (Y_j - \hat{\delta})\} \leq C \tag{2.15}$$

where C is determined such that

$$P(f(Y) \leq C \mid \theta_1 = \theta_2 =, \cdots, = \theta_N) = 1 - \alpha \tag{2.16}$$

After the discussion below, we have the explicit form for $f(Y) = n\bar{Y}_A^T \Sigma_A \bar{Y}_A$, where $A = \{i \mid \bar{Y}_i < 0\}$, and $\bar{Y}_A$ is $\bar{Y}$ with ith element remaining for $i \in A$ and Σ_A is Σ with ith row and columns remaining for $i \in A$.

To find $f(Y)$, let $A \subseteq \{1, 2, \cdots, N\}$ be a subset of indices. And $\bar{Y}^j$ be the jth average of Y (i.e. $\bar{x}_i - \bar{x}_j$). $j = 1, 2, \cdots, N$. It can be calculated that for the 'quadrant' that $\bar{Y}^j$ is smaller than 0 for $j \in A$ and other average less or equal to 0, $f(Y)$ in (2.15) becomes $n\bar{Y}_A^T \Sigma_A \bar{Y}_A$. Where $\bar{Y}_A$ is a vector with length $|A|$ (cardinality of A), and the elements are $\bar{Y}^j$ for $j \in A$. And Σ_A is this sub-matrix of Σ with all jth row and column remaining (still $j \in A$). So our test statistic is $\bar{Y} = (\bar{Y}^1, \cdots, \bar{Y}^N)$, keep in mind that these are differences between average of the ith population and others, so it's $N - 1$ dimension.

The accept region of likelihood ratio test is something between a cube (as in Gupta's case) and an ellipsoid. To illustrate, let's now assume that we have N=3 and we first test whether population 3 should be taken into $\hat{S}$.

Let's draw a 2-d Cartesian coordinate system, with horizontal axis denotes $\hat{\delta}_1 = \hat{\theta}_3 - \hat{\theta}_1$ and vertical axis $\hat{\delta}_2 = \hat{\theta}_3 - \hat{\theta}_2$.

So the using Gupta's rule, if a point lies above $x = -d$ and to the right of $y = -d$ then we accept the null.

For the likelihood ratio test rule, if a point lies in the first quadrant it will be accepted. If it's in the second quadrant, the null is accepted if its first coordinate

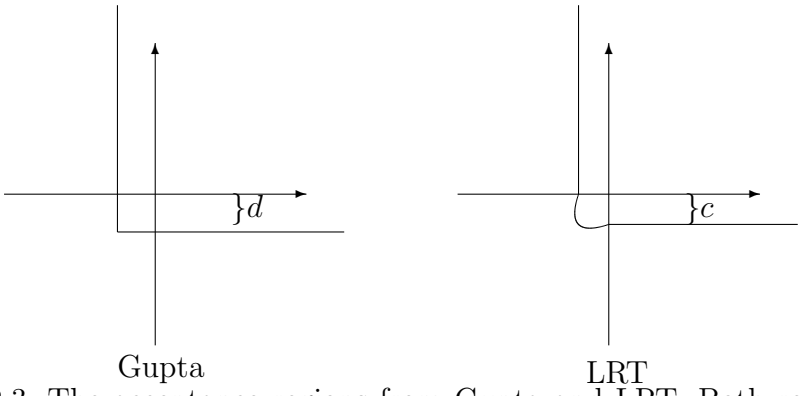

Figure 2.3: The acceptance regions from Gupta and LRT. Both regions are the top right of the two curves.

is larger than $-c$, similarly if it's in the fourth quadrant, the null is accepted if its second coordinate is larger than $-c$. If it's in the third quadrant, it's accepted if it lies to the upper right of the curve $x^2 - xy + y^2 = c^2$. The picture is shown in figure 2.3

Here are a few remarks:

First they both have continuous boundaries, but both have non-differential points. Gupta's happens at the lower left corner, LRT happens at the intersection of boundary and axes. Besides, this confidence set has non flat boundary comparing with Gupta's and Seal's confidence set.

Second, as Gupta's boundary contains point $(-d, -d)$ and LRT's contains $(-c, -c)$. If $c = d$ then Gupta's acceptance region will be contained in LRT's. So we must have $d > c$.(if $d < c$ then Gupta's acceptance region in the second, third and fourth quadrant will all be contained by that of LRT's.) Note that here $\hat{\delta}_1$ and $\hat{\delta}_2$ are positively correlated.

Also we would note that there is a deficiency for LRT that it's not a monotone test. A monotone test means if it accepts H_0 for a sample with average $(\hat{\delta}_1, \hat{\delta}_2)$, it must accept H_0 for all the samples with average $(\tilde{\delta}_1, \tilde{\delta}_2)$ if $\hat{\delta}_1 \leq \tilde{\delta}_1$ and $\hat{\delta}_2 \leq \tilde{\delta}_2$. This can be visually verified from the graph of acceptance regime above.

A second deficiency is that usually we don't know the variance (which is the com-

mon case). Then under the null the δ that maximizes (2.14) is still the same, but the Σ that maximizes (2.14) will be very complicated. To be precise, consider the (i,j)th element of $\hat{\Sigma}$, this is the MLE of covariance of the ith and jth population, will be $\frac{1}{n}\sum_{k=1}^{n}(Y_{ik}-\tilde{\delta}_i)(Y_{jk}-\tilde{\delta}_j)$, where Y_{ik} is the kth observation of the ith population, (similarly for Y_{jk}), and $\tilde{\delta}_i=\hat{\delta}_i 1(\hat{\delta}_i\geq 0)$ where $\hat{\delta}_i$ is the sample average. So plugging $\hat{\Sigma}$ into (2.15) we get a very complicated form to solve.

All we have discussed are methods that can get a confidence set $\hat{S}$ in one step. Another idea is that we can screen out indices one by one until we cannot distinguish them. The method below follows this idea.

For a hypothesis testing, if our null is simple, i.e. null space is a single point, then we can use the p-value method. If the p-value, the probability of 'more extreme' situation happens given the null is true, is below our confidence level α, then we reject the null. One idea is that we sequentially test that the remaining population are equally the best, and kick out a population that 'performs' the worst if the null is rejected, and go on to the next test until we accept that the remaining populations are equally good. This idea was proposed by *Hansen et al.* (2011). Although there are concern that these sequential testings can 'accumulate' Type I errors with unfortunate consequences, in their paper, the authors claimed that the Model Confidence Set (MCS) does not suffer from this problem because the sequential testing is halted when the first hypothesis is accepted. Below is the detailed procedure.

If our set of remaining populations is $M=\{i_1,\cdots,i_m\}$, and the null is that $H_{0,M}: \quad \theta_{i_1}=\cdots=\theta_{i_m}$ with the alternative $H_{1,M}=H_{0,M}^c$. Recall our notation $\delta_{ij}=\theta_i-\theta_j$ and naturally $\hat{\delta}_{ij}=\hat{\theta}_i-\hat{\theta}_j$, we define $d_{i\cdot}=\sum_{j\in M}\delta_{ij}/m$ and naturally $\hat{d}_{i\cdot}=\sum_{j\in M}\hat{\delta}_{ij}/m$. From these statistics, we construct the t-statistics

$$t_{ij}=\frac{\hat{d}_{ij}}{\sqrt{\widehat{\mathrm{VAR}}(\hat{d}_{ij})}} \quad \text{and} \quad t_{i\cdot}=\frac{\hat{d}_{i\cdot}}{\sqrt{\widehat{\mathrm{VAR}}(\hat{d}_{i\cdot})}} \tag{2.17}$$

where $\widehat{\mathrm{VAR}}(\hat{d}_{ij})$ and $\widehat{\mathrm{VAR}}(\hat{d}_{i\cdot})$ denote estimates of $\mathrm{VAR}(\hat{d}_{ij})$ and $\mathrm{VAR}(\hat{d}_{i\cdot})$, respectively. The t-statistics t_{ij} and $t_{i\cdot}$ are associated with the null hypothesis that H_{ij} : $\delta_{ij} = 0$ and $H_{i\cdot}$: $\delta_{i\cdot} = 0$ because we have

$$\begin{aligned} \theta_{i_1} = \cdots \theta_{i_m} &\Leftrightarrow \delta_{ij} = 0 \text{ for all } i, j \in M \\ &\Leftrightarrow \delta_{i\cdot} = 0 \text{ for all } i \in M \end{aligned}$$

Moreover, the equivalence extends to $\{\delta_{i\cdot} \leq 0 \text{ for all } i \in M\}$ as well as $\{|\delta_{ij}| \leq 0 \text{ for all } i, j \in M\}$, and these two formulations of the null hypothesis map naturally into the test statistics

$$T_{\max,M} = \max_{i \in M} t_{i\cdot} \quad \text{and} \quad T_{R,M} = \max_{i,j \in M} |t_{ij}|$$

which are available to test the hypothesis $H_{0,M}$. Both tests reject $H_{0,M}$ if the test statistics are above a certain number. The asymptotic distributions of these test statistics are nonstandard because they depend on nuisance parameters (under the null and the alternative). However the nuisance parameters pose few obstacles, as the relevant distributions can be estimated with bootstrap methods that implicitly deal with the nuisance parameter problems.

Once we reject a null $H_{0,M}$, we need an elimination rule e_M to delete one population from M. For these two tests, the author proposed two natural elimination rules $e_{\max,M} = \arg\max_{i \in M} t_{i\cdot}$ and $e_{R,M} = \arg\max_{i \in M} \sup_{j \in M} t_{ij}$. They proved that these two elimination rules equipped with the testing rules $T_{\max,M}$ and $T_{R,M}$ satisfy the assumption that under the alternative, as sample size n goes to infinity, the probability of deleting (possibly one of) the best population goes to zero.

2.2.3 Bayesian approach

Bayesian approach is another useful tool. *Roth* (1978) proposed a method called fiducial procedure. By considering θs as random variables instead of fixed unknown numbers, he calculated p_i, the probability of the ith population being the best based on the samples. Then his selection rule is: arrange $\{p_i\}$ in descent order $p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(N)}$, then sequentially select $(N), (N-1) \cdots$ until (i), into $\hat{S_F}$, where i is the largest number such that $\sum_{j=i}^{N} p_{(j)} \geq 1-\alpha$. In the calculation of p_i, he assumed that all θs are normally distributed with means equal to the corresponding sample average and a common known variance. In his paper, he proved that first, for any observation, his confidence set is a subset of that from Gupta's rule. Second, if there are more than two populations, then there is positive probability that fiducial confidence set is a proper subset of Gupta's. But smaller size must suffer from some deficiency. In his paper he also showed that the fiducial procedure does not satisfy the condition $\inf_\theta P(\theta_{(N)} \in \hat{S_F}) \geq 1-\alpha$ when there are more than two populations, the worst case is just like Gupta's, i.e. $\theta_{(1)} = \theta_{(2)} = \cdots = \theta_{(N-1)} = \theta_{(N)} - \epsilon$ and $\epsilon \to 0$.

But as we are in the Bayesian world, maybe this frequentist's P^* condition (i.e., contain the fixed "true" best index with a desired probability) is not appropriate for a Bayesian method. *Gupta and Yang* (1985) extended Roth's work. They proposed a non-randomized method ψ^B and a randomized method ψ^{B*}.(We will see the fiducial procedure is a special case of ψ^B if we choose non-informative prior for the mean vector). We introduce these two methods below:

Let X_{ij} be the jth observation of the ith population. $i = 1, \cdots, N$, $j = 1, \cdots, n_i$. Let $X_i = T_i(X_{i1}, \cdots, X_{in_i})$ be a suitable estimator of θ_i (usually the sample average). A selection rule based on an observation $X = (X_1, \cdots, X_N)$ will be denoted by $\psi(x) = (\psi_1(x), \cdots, \psi_N(x))$ where $\psi_i(x) : \mathbb{R}^N \to [0,1]$ is the probability that i is included in the confidence set when $X = x = (x_1, \cdots, x_N)$ is observed. A correct selection (CS) is defined to be the selection of any subset that includes the best

population. Suppose we have a prior distribution for θ and instead of satisfying the frequentist's P^* condition (defined last paragraph), they considered controling the posterior probability of CS to be larger than $P^* = 1 - \alpha$, that is

$$P(CS|\psi, X = x) = \sum_{i=1}^{N} \psi_i(x) p_i(x) \geq P^* \quad \forall x \tag{2.18}$$

where $p_i = P(\theta_i \geq \theta_j \ \forall j | X = x)$. If we assume that the posterior cdf of θ is continuous then it's clear that $\sum_{i=1}^{N} p_i(x) = 1 \ \forall x$. Let $p_{[1]}(x) \leq \cdots \leq p_{[N]}(x)$ be the ordered $p_i(x)$'s and let $\theta_{(i)}$ be the population associated with $p_{[i]}(x)$ (Note that here $\theta_{(i)}$ no longer denote the ith smallest population mean as now they are not fixed). Then a subset selection procedure is completely specified by $\{\psi_{(1)}(x), \cdots, \psi_{(N)}(x)\}$ where

$$\psi_{(i)}(x) = P((i) \text{ is selected} \mid \psi, X = x), \quad i = 1, \cdots, N \tag{2.19}$$

The posterior-P^*-condition (called PP^*-condition) is defined to satisfy the inequality in (2.18) as well as $\psi_{(N)}(x) = 1 \ \forall x$ (means we always select the population with highest probability being the best based on the observation). Let's have a look at (2.18), it means *on average* (w.r.t the posterior distribution), that this selection has good behavior (in the sense of selecting the true best population with high probability), which is weaker than the 'level $1-\alpha$' requirement since this requires *everywhere* good behavior.

Now define non-randomized rule $\psi^B = (\psi^B_{(1)}(x), \cdots, \psi^B_{(N)}(x))$ by:

$$\psi^B_{(i)}(x) = \begin{cases} 1 & \text{if } i \geq j \\ 0, & \text{if } i < j \end{cases} \tag{2.20}$$

and j is the largest integer between 1 and N such that

$$\sum_{i=j}^{N} p_{[i]}(x) \geq P^*$$

So we can see this is very like the fiducial procedure's idea. And if we allow random selection, a similar rule $\psi^{B*} = (\psi^{B*}_{(1)}(x), \cdots, \psi^{B*}_{(N)}(x))$ can be formed by setting $\psi^{B*}_{(N)}(x) = 1$ and

$$\psi^{B*}_{(i)}(x) = \begin{cases} 1 & \text{if } \sum_{j=i}^{k} p_{[j]}(x) \leq P^*,\ i \neq k, \\ \nu & \text{if } \sum_{j=i+1}^{k} p_{[j]}(x) < P^* \text{ and } \sum_{j=i}^{k} p_{[j]}(x) > P^* \\ 0 & \text{otherwise} \end{cases} \tag{2.21}$$

where the constrain ν is determined such that

$$\nu p_{[i]}(x) + \sum_{j=i+1}^{k} p_{[j]}(x) = P^* \quad 0 < \nu < 1$$

In this paper, the authors again considered that given the θs, the observations of population i has $iid\ N(\theta_i, \sigma_i^2)$ distribution.

First they consider non-informative prior $\tau(\theta) \propto C$ with common known σ^2 and common sample size n, then

$$p_i(x) = \int_{-\infty}^{\infty} \prod_{j \neq i} \Phi(t + \sqrt{n}\sigma^{-1}(x_i - x_j)) \mathrm{d}\Phi(t), \quad i = 1, \cdots, N$$

which coincides the fiducial procedure.

If they have unequal but known σ_i^2 and unequal sample sizes n_i, then

$$p_i(x) = \int_{-\infty}^{\infty} \prod_{j \neq i} \Phi(t\nu_i/\nu_j + (x_i - x_j)/\nu_j) \mathrm{d}\Phi(t), \quad i = 1, \cdots, N$$

where $\nu_i = \sigma_i/\sqrt{n}$.

If they have unequal and unknown σ_i^2 and unequal sample sizes n_i, they chose prior $\tau(\theta_i, \sigma_i) \propto \sigma_i^{-1}$ for each population and

$$p_i(x) = \int_{-\infty}^{\infty} \prod_{j \neq i} T_{\nu_j}\Big(t\frac{s_i/\sqrt{n_i}}{s_j/\sqrt{n_j}} + \frac{x_i - x_j}{s_j/\sqrt{n_j}}\Big) \mathrm{d}T_{\nu_j}(t)$$

where $\nu_i = n_i - 1$, $s_i^2 = \sum_{j=1}^{n_i}(x_{ij} - x_i)^2/\nu_i$, $i = 1, \cdots, N$, and T_ν is the *cdf* of the t-distribution with ν degrees of freedom. For large ν, it can be approximated by the normal distribution.

They also considered two more cases, one is that θs have common distribution $N(\mu_0, \sigma_0^2)$ and given θ_i, X_i (sample average) has distribution $N(\theta_i, \sigma^2/n)$,then

$$p_i(x) = \int_{-\infty}^{\infty} \prod_{j \neq i} \Phi(t + bn\sigma^{-2}(x_i - x_j)) \mathrm{d}\Phi(t), \quad i = 1, \cdots, N$$

where $b^2 = (\sigma_0^{-2} + n\sigma^{-2})^{-1}$.

The other one is θ_i have independent normal prior distribution $N(\mu_i, \sigma_{0i}^2)$. And given θ_i, X_i has distribution $N(\theta_i, \sigma_{1i}^2/n_i)$, let $z(x_i) = b_i^2(\sigma_{0i}^{-2}\mu_i + n_i\sigma_{1i}^{-2}x_i)$, $b_i^2 = (\sigma_{0i}^{-2} + \sigma_{1i}^{-2}n_i)^{-1}$, and

$$p_i(x) = \int_{-\infty}^{\infty} \prod_{j \neq i} \Phi(tb_i + (z(x_i) - z(x_j))/b_j) \mathrm{d}\Phi(t), \quad i = 1, \cdots, N$$

Comparing with the frequentists' methods, this Bayesian approach has some advantages. As listed in *Gupta and Yang* (1985):

a) It can be applied to any family of distributions, even their mixture, and does not need equal sample sizes.

b) Good prior information will not be ignored. Even under non-informative situation they perform well.

c) They are robust in terms of the loss function. We do not even need to specify or to know the exact form of the loss function before we make a decision. As long as the loss function has some 'regular' property like transition invariant and non-increasing in the size of confidence set.

d) The weight or contribution of each population in the confidence set is known.

e) It is robust if the true family of distributions for each population is symmetric.

2.2.4 Comparison of these methods

Evaluation of confidence set is closely related to the measure (in most cases Lebesgue) of the confidence region of parameters (in our case, is the mean vector), because it's converted from a confidence region. *Joshi* (1969) and *Hwang and Casella* (1982) considered the confidence region that minimizes the maximum Lebesgue measure. Basically their method is based on a ball confidence region $\{\mu \mid \| \bar{X} - \mu \|_2^2 \leq C\}$, and change $\bar{X}$ to $\delta(\bar{X})$ to achieve the minimax property. *Hooper* (1982) considers a region based on a Neyman-Pearson type criterion to form confidence region (dependent on the true mean vector), which minimizes the expected Lebesgue measure among invariant confidence sets.(A confidence region $C(x)$ with data x generated from model with parameter θ is invariant if the conditional frequency (given $X = x$) of $C(x)$ covering θ is the same as that of $C(g(x))$ covering $g(\theta)$ and g belongs to a invariant group, usually the group of linear transformation).

Cohen et al. (1973) proposed a criterion of judging the behavior of a confidence region. A confidence region is said to be admissible if there is no other confidence set that has smaller Lebesgue measure and probability of covering the true parameter.

For one dimension confidence interval, *Aitchison and Dunsmore* (1968) and *Winkler* (1972) considered the loss function which is a linear combination of undershooting (upperbound of confidence interval smaller than the estimated parameter) and overshooting (lower bound larger than the parameter) and the length of the interval. They

showed that if the penalty for the length is too large then the confidence interval will shrink to a point estimator. For multi-dimension, *Casella and Hwang* (1991) used the loss function which is the linear combination of the Lebesgue measure of the confidence set and the probability not covering the true parameter. They didn't directly constrain the probability of covering the true parameter, but it is connected with the penalty of not covering it, i.e the penalty parameter of not covering the truth is a one-to-one function of the constrain of the probability of covering the truth.

But smaller confidence region for the mean does not necessarily implies smaller expected confidence set in our setting. (When we talk about the measure of a set , we mean the number of indices it contains). We can see that Gupta's confidence region of means is 'one-sided', i.e. have infinite Lebesgue measure, but contains fewer indices comparing to the minimum confidence region (The ellipsoid).

Lehmann (1952) considered maximizing the minimum power. In this article he proved (Thm 4.1) that in our setting, in three population case (which means every test involves two-dimension parameter testing), the Gupta type acceptance region maximizes the minimum power. The minimum power always happens where the true (δ_1, δ_2) lies in either $\{0, \infty\}$ or $\{\infty, 0\}$, this corresponds to the case that two of the three population have very large means comparing to the third, while the difference between these two large means is small. Intuitively this makes sense, as in this scenario the question is actually the comparison of only two populations (since the third one has such small mean), but the critical value d is unnecessarily enlarged and thus has small power (we can see that d gets larger as the dimensionality gets larger from the way d in constructed). But here we should know the variance among all the monotone hypothesis.

One concern is that maximizing the minimum power might not be the best choice. Since what we would like is the 'good' behavior around the true parameters, not the parameters far away (i.e. not likely to be true given the samples).

Another kind of loss function is instead of the measure of the confidence region of estimation of θ, we directly penalize the size of the confidence set. *Bjørnstad* (1986) considered the following two kinds of loss functions

$$l(\theta, \hat{S}) = \alpha(|\hat{S}|) \sum_{i \in \hat{S}} l^0(\theta_{(N)} - \theta_i) \quad l(\theta, \hat{S}) = \alpha(|\hat{S}|) \sum_{i \in \hat{S}} l^0(\theta_{(N)}/\theta_i - 1) \tag{2.22}$$

Here $|\hat{S}|$ means the size of $\hat{S}$. α, l^0 are continuous, non-decreasing and $l^0(0) = 0$ and $l^0(x) > 0$ if $x > 0$.

To understand this, first the loss is positive related with the size of the confidence set, then we penalize more if the true best mean is away from others. In his paper the author discussed the asymptotic behavior of Gupta's rule. He proved that as the sample size n goes to infinity, if there is only one best population (treated as fixed, i.e. not Bayesian view), then Gupta's method will eventually select the only true population, i.e. the false positive and false negative rate will both go to zero. Further more, this convergence holds uniformly for a set of θ lies in a compact set $\mathcal{K}$, i.e. if we restrict $\theta \in \mathcal{K}$, then for any $\epsilon > 0$, there exists an N such that for any sample size $n > N$ we have the probability of including the best and not including others both larger than $1 - \epsilon$, regardless of where exactly θ lies, (as long as it belongs to $\mathcal{K}$). This seems too greedy since what we want is just covering the truth with probability $1 - \alpha$. But here if we use hypothesis testing point of view, our null hypothesis is that a vector (differences between the mean of the ith population and that of the others) lies in a set that has interior points. So in order to have good coverage for all points in the null space, like the argument in deriving (2.3), it's equivalent to cover the 'sharp point' (the origin) when the true vector lies there. So if the truth is that the parameter lies inside the null space, by law of large number the sample mean converges uniformly to the true population mean, thus we have that the sample difference vector of the largest population will eventually fall into the null space (thus selected into the con-

fidence set) and sample difference vector of other population will converge to points that are away from the null space, thus get excluded in the confidence set eventually.

If our main concern is just the expected size of our sample size $E_\theta(|\hat{S}_\psi|)$, (here the subscript θ of E means the expected size depends on the true mean vector θ, and the subscript ψ of $\hat{S}$ means the confidence set $\hat{S}$ follows the selection rule ψ) *Berger et al.* (1979) considered minimizing the largest possible expected size, i.e. minimizing $\max_\theta E_\theta(|\hat{S}_\psi|)$. In his paper, he proved that Gupta's method achieved the lower bound, which is $(1-\alpha)N$, which happens at the case one of the population is slightly larger than each of the rest.

Recall that Gupta's test (in fact the whole UIP test) is monotone while likelihood ratio test is not. *Gupta and Huang* (1980) showed that for any non-monotone test, we can construct a better new test based on it, in terms of smaller maximum expected set size (i.e. how many elements are there in a set), while still preserving the level $1-\alpha$.

But by just looking at the maximum, we sacrifice a lot. A method having a smaller maximum size might perform relatively worse in general, comparing with a method having a larger maximum size. What's more, if we just consider the maximum expected size, then Gupta's method is only as good as the naive selection rule, which selects any population with probability $1-\alpha$ and obviously satisfies selecting the best population with probability no less than $1-\alpha$ for any true θ. And $E_\theta|\hat{S}| = \sum_{i=1}^N P(i \in \hat{S}) = \sum_{i=1}^N (1-\alpha) = (1-\alpha)N$ for any θ, so $\max_\theta E_\theta|\hat{S}| = (1-\alpha)N$. Thus this method also achieves the lower bound of maximum expected set size. Even if the performances are the same for these two methods at the point where all populations have the same mean (in this case the best population is tagged to be any one of them), Gupta's method selects much fewer in many other cases. So instead of just looking at the maximum expected size, maybe it's a good idea to look at the general behavior,

i.e. some sort of average expected size with respect to some prior distribution of θ, or just $\int_\Omega E_\theta(|\hat{S}_\psi|)\mathrm{d}\theta$ if we know $\theta \in \Omega$ a compact subspace of R^N.

Following the idea of *Lehmann* (1952), we want to find $\{d_{ij}\}$ satisfying (2.7) and have some good property. Another criterion is instead of maximizing the minimum power, we try to maximize the average power, here average is taken w.r.t some distribution of the true parameter. We want to give more reward of enlarging the power of those parameter around the point estimator of it since they are more likely to be in that area given the data. So we can use the posterior distribution of the mean vector (we can use non-informative improper prior). Thus, our test statistic will be something other than the differences of sample average (plus estimator of variance if it's not known).

Next we consider loss functions in Bayes world.

The PP^* condition, which is defined after formula (2.19), is:

$$P(CS|\psi, X = x) \geq P^* \quad \forall x \text{ and } \psi_{(N)}(x) = 1 \tag{2.23}$$

And for given a prior τ, let $D = D(\tau, P^*)(D^* = D^*(\tau, P^*))$ be the class of all non-randomized (randomized) selection procedures in which all procedures satisfy the PP^* condition for any given observation $X = x$. We can see that D and D^* contain many rules including some useless rules like $psi_i(x) = 1$ for all i and x (i.e. select all population regardless of the observation). They all satisfy the PP^* condition hence are valid candidates. Our next job is to select some rule from them to have good properties in some sense.

Another definition is for a group G (usually permutation), for all $g \in G$,$\theta \in \Theta$ and confidence set s, a loss function L has property T if (1) $L(\theta, s) = L(g\theta, gs)$, (2) $L(\theta, s)$ is non-increasing in θ_i for $i \in s$ and (3) $L(\theta, s) \leq L(\theta, s')$ if $s \subseteq s'$. Basically the first property requests the selection is invariant under permutation, second is we

don't penalize more if the true population mean get higher, when this population index is in the confidence set, and the third means you always penalize more for a larger confidence set. So if we define $L(\theta, s) = |s|$ where $|s|$ is the size of s, then this loss function has property T. Then *Gupta and Yang* (1985) proved the following theorem:

Suppose the prior distribution τ is symmetric on Θ. Given $\theta \in \Theta$, $X_1, \cdots, X_N$ are independently distributed and the pdf $f(x_i|\theta_i)$ has monotone likelihood ratio property. Then the selection procedure $\psi^B(\psi^{B*})$ defined in (2.20) and (2.21) is monotone and is a Bayes procedure in $D(D^*)$ provided that the loss function has property T. This theorem also holds if $H(\theta|x)$, the posterior cdf of θ, given $X = x$, is absolutely continuous and have the generalized stochastic increasing property, that is: (1) $H(\theta|x) = \Pi_{i=1}^N H_i(\theta_i|x)$, where $H_i(\cdot|x)$ is the posterior cdf of θ_i and (2) $H_i(t|x) \geq H_j(t|x)$ for any t whenever $x_j \leq x_j$. To understand (2), let's assume x_i is the sample average of the ith population having $N(\theta_i, \sigma^2/n)$ and we use non-informative prior, x_is independent given θ_is, so $H_i(t|x) = H_i(t|x_i) = \Psi(t\sqrt{n}/\sigma - x_i)$ so it's non-increasing in x_i.

These two ψ methods are also the most efficient in $D(D^*)$, i.e.eff$(\psi) \geq$ eff(ψ') for all $\psi' \in D(D^*)$. The efficiency is defined as

$$\text{eff}(\psi|x) = P(CS|\psi, x)/E(S|\psi, x)$$

where $E(S|\psi, x)$ is the posterior expected size of the selected subset. The expectation of eff$(\psi|x)$ is the efficiency of procedure and is denoted by eff(ψ).

In summary, there are good properties we want a selection procedure to have. Below, a procedure is denoted as $\psi(\hat{\theta}_1, \cdots, \hat{\theta}_N, \hat{\Sigma}_N, n)$, where $\hat{\theta}_i$ are consistent estimator of the mean of each population (in most cases sample mean) and $\hat{\Sigma}_N$ is a consistent

estimator of the covariance matrix of the estimator of vector θ, (pooled sample covariance matrix in general), and n the common sample size of each population. In other words, as all the popular methods (Seal, Gupta, UIP, LRT, all the Bayesian methods) concern only the sample averages (actually only the differences of those averages if sample sizes of each populations are the same.), and the estimated covariance matrix (if unknown), so we only focus on selection rules that are function of only $\hat{\theta}$ and $\hat{\Sigma}_N$. This is due to the reduction to sufficient statistics for normal distribution.

Below are some good properties we would like a selection rule to have: (denote that event that the selection rule selects the true best population as correct selection, or CS, and the size of a set $\hat{S}$ which counts how many populations it contains by $|\hat{S}|$, a set followed a rule ψ is denoted as $\hat{S}_\psi$)

I Include the best population with high probability, regardless of where the true mean vector lies (frequentists' view), i.e. $\inf_\theta P(CS|\psi) \geq 1-\alpha$.

I'. Include the best population with high probability on average (Bayesian view), i.e. $EP(CS|\psi) \geq 1-\alpha$, where expectation is taken with respected to some posterior distribution of θ.

II. Small maximum expected set size, i.e. small $\max_\theta E_\theta|\hat{S}_\psi|$, where expectation is taken with the distribution of $\hat{\theta}$ for fixed θ.

II'. Small average expected set size, i.e. small $E_\tau E_\theta|\hat{S}_\psi|$ where τ is some posterior distribution of θ.

III. Monotonicity, meaning if population i is selected into $\hat{S}$ by some rule ψ based on $\hat{\theta}$, then increasing $\hat{\theta}_i$ while remains other $\hat{\theta}_j$ the same will not exclude i, follow the

same ψ. i.e. $1(i \in \hat{S}|\psi(\hat{\theta}_1, \cdots, \hat{\theta}_i, \cdots, \hat{\theta}_N) \geq 1(i \in \hat{S}|\psi(\hat{\theta}_1, \cdots, \hat{\theta}'_i, \cdots, \hat{\theta}_N)$, for any $\hat{\theta}_i \geq \hat{\theta}'_i$.

IV. If we add a 'useless' competitor, it shouldn't effect the choice of S_ψ, i.e. $\lim_{\hat{\theta}_N \to -\infty} S_{\psi(\hat{\theta}_1, \cdots, \hat{\theta}_N, \hat{\Sigma}_N, n)} = S_{\psi(\hat{\theta}_1, \cdots, \hat{\theta}_{N-1}, \hat{\Sigma}_{N-1}, n)}$, where $\hat{\Sigma}_{N-1}$ is the first $N-1$ columns and rows of $\hat{\Sigma}_N$.

V. As the differences of $\{\hat{\theta}_i\}$s get smaller and smaller, we tend to select all of them. i.e. let $\Delta = \max_{i,j}\{|\hat{\theta}_i - \hat{\theta}_j|\}$, we want $\lim_{\Delta \to 0} \hat{S}_\psi = \{1, 2, \cdots, N\}$.

VI. If a population 'performs' the best, it should always be selected into the confidence set. i.e. $\hat{\theta}_i \geq \hat{\theta}_j \ \forall j \Rightarrow i \in \hat{S}_\psi$.

Here are a few remarks:

0. Here we mainly compare three methods: Gupta's, model confidence set (MCS) and Bayesian (which is Fiducial if we consider non-informative prior). We don't consider much of LTR method because in the situation we don't know the variance (which is almost all the cases), it's very hard to get the form of the test.

1. I is much stronger than I ' and can imply I '. But as the 'inf' always takes place where all θ_is are equal. If the sample means are vary dispersed, then the truth being all means are the same is vary small. Gupta's method achieves I, but sacrifices a lot in the sense of large set size. In order words, Gupta's method always considers the worse case, even if it seems very unlikely to be the case, and thus unnecessarily enlarges the set size. For example, if our observations are $(10, 9.9, 9.3, 9.3, 0, 0)$, with common sample size $n = 18$, known standard deviation $\sigma = 1$ and significance level $\alpha = 0.05$. As here our total population is $N = 6$ so $d = 2.234$ and $d\sigma\sqrt{2/n} = 0.745$ so we should include the first four populations into our $\hat{S}$. But if we think that as the last two

means are more than 10 standard errors away from the largest sample average, it's almost impossible for these two populations to be the best (p-value almost zero), so there are 'essentially ' just four populations competing, we use $N = 4$ hence $d = 2.062$ and $d\sigma\sqrt{2/n} = 0.687$, so we only select the first two populations into our $\hat{S}$. This idea will be applied in the nest section where we will introduce two new methods call step-up and step-down methods.

2. Bayesian method doesn't satisfy I when $N > 2$ but satisfies I'.

3. MCS satisfies more than I. To see this, first we consider that all populations have equal means, then what we concern is the probability of including the first population in our $\hat{S}$. This probability is equal to the probability of accepting the null which is all population means are equal (in this situation our $\hat{S}$ includes all indices), plus the probability of rejecting the null but finally didn't kick out this population in the following steps. The first probability is already $1 - \alpha$, and the second is strictly positive (for example only one other population's sample mean is way below others, we will kick this one out and terminate). So in summary the probability of having the first index in $\hat{S}$ is larger than $1 - \alpha$ when all population means are equal. And if it is not the case, the one with largest mean will be harder to kick out than the case when all means are equal.

4. All three methods satisfy III (but LRT doesn't).

5. Gupta's method violates IV, because each new-comer will enlarge the 'selection threshold' d, which depends only on number of populations. But Bayesian method satisfies this property. Because as an added population's sample mean goes to negative infinity, its probability of being the best will shrink to zero and won't effect the probability of others being the best. MCS also satisfies this. If we add a useless competitor, it will be kicked out in the first step. More over, MCS does a better job in the sense of avoiding the influence of bad population. Because as long as the sample mean of a bad population is below a certain threshold, it will completely have

no impact on the selection procedure. But for the other two, the influence of bad population always exists (becomes smaller and smaller as the sample mean of the bad population gets lower).

6. Gupta's method and MCS satisfy V, it's very easy to see this from the rule. Bayesian method doesn't always satisfy this, but only when the population is large (larger than $1/\alpha$). For example, if $\alpha = 0.1$, then we only have a problem with V when there are at least 10 populations.

7. All the methods introduced in this review satisfy VI.

So in summary, Gupta's method satisfies I,III,V and VI, and Bayesian I',III,IV,VI, and MCS I,III,IV,V,VI. In the consideration of II and II' Gupta's method focus on the worst case scenario, which is all the true means are equal, so it performs well when this is the truth, or the differences between the true population means are small. But it performs bad when the true differences of population means are large, in which case Bayesian method performs better, because Gupta's method's lack of ability to get rid of 'clearly' bad interrupting population. And MCS performs worst, as it selects every population with larger probability (which is discussed in remark 3).

Gupta's method performs better if the true means are close and Bayesian method is better when true means are separated. So a natural question is that can we have a method performs better for both cases, and satisfies as many above properties as possible.

2.3 Step-up and step-down methods

2.3.1 Introduction

As has been discussed in the last section, Gupta's method always considers the worst case scenario, i.e., when all the population means are the same, even if there are

strong evidence that some of the population means are not the best. By considering the worst case, the "threshold" d is unnecessarily large and thus to not have enough power to exclude true inferior indices. Following the sequential testing idea from MCS, we propose two multi-stage procedure of constructing the confidence set $\hat{S}$. Section 2.3.2 will contain the introduction of the two methods, followed by a numerical study in section 2.3.3. In section 2.3.4 we will have a discussion on these two methods.

2.3.2 Form of the two methods

Before the introduction of the two methods, we first review the procedure of Gupta's method.

We have a critical number d_N such that d_N satisfies

$$P(Z_i \leq d_N \ \ \forall i = 1, \cdots, N-1) = 1 - \alpha \tag{2.24}$$

Where $Z = (Z_1, \cdots, Z_{N-1})$ follows a multivariate normal distribution with mean zero, its variance matrix has 1 on its diagonal and 0.5 elsewhere. So Gupta's selection rule is

$$i \Rightarrow \hat{S}_G \quad iff \quad \hat{\theta}_i \geq \hat{\theta}_{(N)} - \frac{\sqrt{2} d_N \sigma}{\sqrt{n}} \tag{2.25}$$

Where $\hat{\theta}_{(N)}$ is the value of the largest sample means.

Now we introduce the procedures of the two new methods. Note that one key assumption is that they both need that the variance of each point estimators of the population means are the same. For the introduction of the two procedures, we assume for simplicity that we know this common variance and denote it as σ^2. We can also have a similar procedure by using $\hat{\sigma}^2$, the consistent estimator of σ^2. Besides, they also both required that these estimators are independent.

Step-up method:

We have a sequence of critical values $c_1, \cdots, c_N, \cdots$. We sort the value of sample means in ascending order: $\hat{\theta}_{(1)} < \cdots < \hat{\theta}_{(N)}$. (As we have continuous distribution so the chance to tie is zero). And below is the selection rule of step-up method:

Step 0: Initialize an empty set denoted as $\hat{S}_{SU}$.

Step 1: (N) always in $\hat{S}_{SU}$.

Step 2: If $\hat{\theta}_{(N-1)} \geq \hat{\theta}_{(N)} - c_2\sigma/\sqrt{n}$ then $(N-1)$ is selected into $\hat{S}_{SU}$ and go to step 3; Else, stop.

Step 3: If $\hat{\theta}_{(N-2)} \geq \hat{\theta}_{(N)} - c_3\sigma/\sqrt{n}$ then $(N-2)$ is selected into $\hat{S}_{SU}$ and go to step 4; Else, stop.

.....

Step $N-1$: If $\hat{\theta}_{(2)} \geq \hat{\theta}_{(N)} - c_{N-1}\sigma/\sqrt{n}$ then (2) is selected into $\hat{S}_{SU}$ and go to step N; Else, stop.

Step N: If $\hat{\theta}_{(1)} \geq \hat{\theta}_{(N)} - c_N\sigma/\sqrt{n}$ then (1) is selected into $\hat{S}_{SU}$. Stop.

This method can be thought as a sequence of hypothesis testing. First we test whether the population with the second largest estimated mean has its population mean no smaller than that of the population with the largest estimated mean, and continue testing until we reject a test (the null will always be that the target population's mean is the largest) or we select all populations.

Now is the discussion of one approach to get the critical values.

For c_2, we need that $P_{\theta_1=\theta_2}(1 \in \hat{S}_{SU}) = 1-\alpha$. After finding c_2 we find c_3 by solving $P_{\theta_1=\theta_2=\theta_3}(1 \in \hat{S}_{SU}) = 1-\alpha$ (*) Here the only unknown number is c_3, the left

hand side of (*) is monotone with respect to c_3 and is below and above $1-\alpha$ when c_3 is zero and infinity. So we can uniquely get c_3. For greater N, c_N will be find similarly.

Step-down method:

We have a sequence of critical values $c_1, \cdots, c_N, \cdots$ (These are different from the cs in step-up method. For comparison of the methods, we will denote the two vectors of cs as c_{su} and c_{sd}) . We sort the value of sample means in ascending order: $\hat{\theta}_{(1)} < \cdots < \hat{\theta}_{(N)}$. (As we have continuous distribution so the chance to tie is zero). And below is the selection rule of step-down method:

Step 1: Select all indices into a set denoted as $\hat{S}_{SD}$.

Step 2: If $\hat{\theta}_{(1)} \leq \hat{\theta}_{(N)} - c_N\sigma/\sqrt{n}$ then (1) is removed from $\hat{S}_{SD}$ and go to step 3; Else, stop.

Step 3: If $\hat{\theta}_{(2)} \leq \hat{\theta}_{(N)} - c_{N-1}\sigma/\sqrt{n}$ then (2) is removed from $\hat{S}_{SD}$ and go to step 4; Else, stop.

.....

Step $N-1$: If $\hat{\theta}_{(N-2)} \leq \hat{\theta}_{(N)} - c_3\sigma/\sqrt{n}$ then $(N-2)$ is removed from $\hat{S}_{SD}$ and go to step N; Else, stop.

Step N: If $\hat{\theta}_{(N-1)} \geq \hat{\theta}_{(N)} - c_2\sigma/\sqrt{n}$ then $(N-1)$ is removed from $\hat{S}_{SD}$. Stop.

This method can be thought as a sequence of hypothesis testing. First we test whether the population with the smallest estimated mean has the largest population mean, and continue testing until we accept a test (the null will always be the target

population has the largest mean) or we removing all populations except the one with the largest sample mean.

Now is the discussion of one approach to get the critical values.

For c_2, we need that $P_{\theta_1=\theta_2}(1 \in \hat{S}_{SD}) = 1 - \alpha$. After finding c_2 we find c_3 by solving $P_{\theta_1=\theta_2=\theta_3}(1 \in \hat{S}_{SD}) = 1 - \alpha$ (*) Here the only unknown number is c_3, the left hand side of (*) is monotone with respect to c_3 and is below and above $1 - \alpha$ when c_3 is zero and infinity. So we can uniquely get c_3. For greater N, c_N will be find similarly.

2.3.3 Simulation Study

As currently we do not have any theoretical results of these two methods. We would like to show their performance by numerical simulation. As only the differences between population means are what we concern, without loss of generality we assume that the mean of the third population is always zero. The mean of the first and second population ranged from 0 to 4.5, common sample size $n = 4$ (As we already know the common σ so we do not require a large sample size) and common known variance $\sigma^2 = 1$. For any point, we generate data 1000000 times and use the three methods to get selected set. Figure 2.4 through 2.6 the difference in average set size between the three methods.

We can see that these two methods both beat Gupta's method, in the sense that they both have smaller expected set sizes than Gupta's. And the difference between these two are not significant.

2.3.4 Discussion

The reason we do not further investigate these two methods is that although they both beat Gupta's method, in the sense that their expected set sizes are always smaller than that of Gupta no matter what the true population mean θ is, the improvement

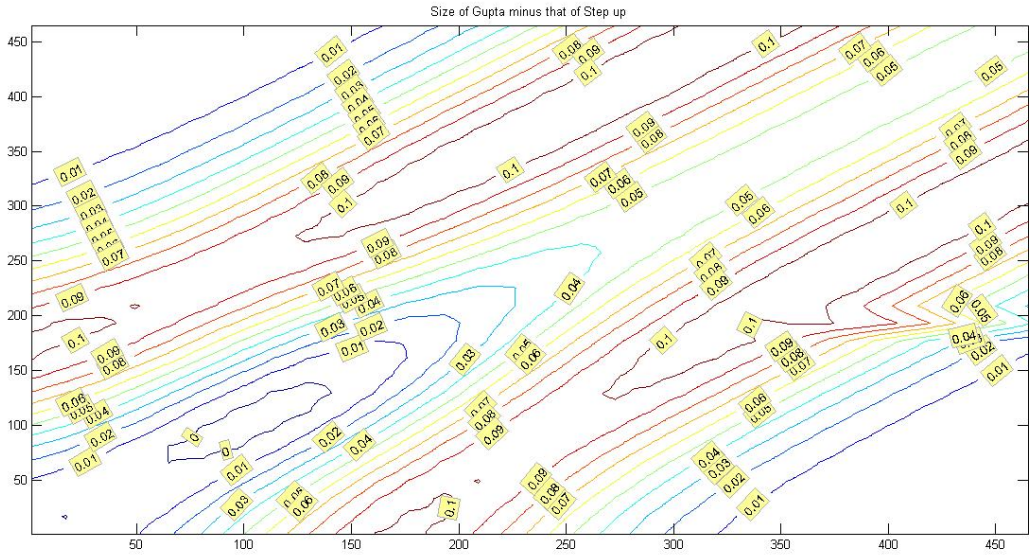

Figure 2.4: The difference of expected set sizes between set from Gupta's method and the step-up method.

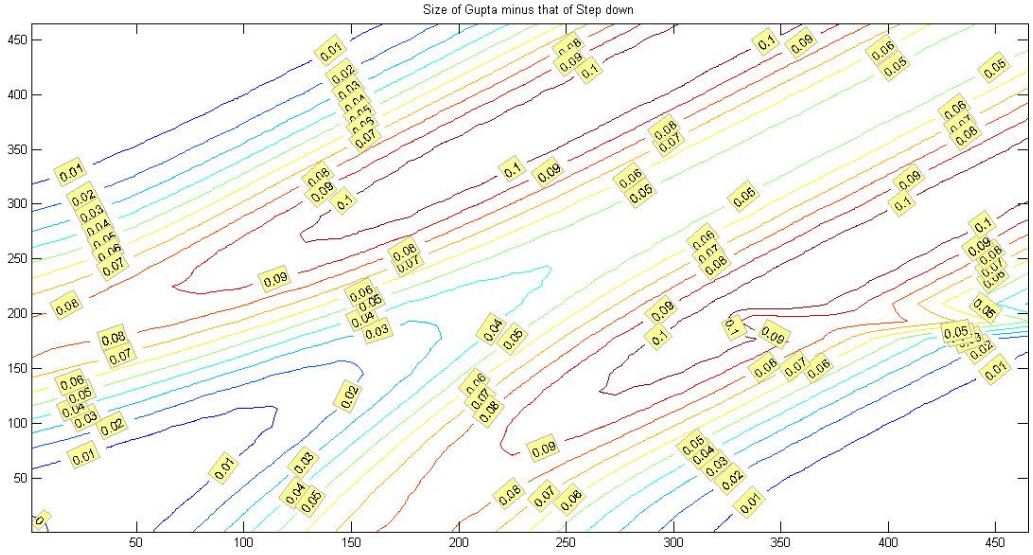

Figure 2.5: The difference of expected set sizes between set from Gupta's method and the step-down method.

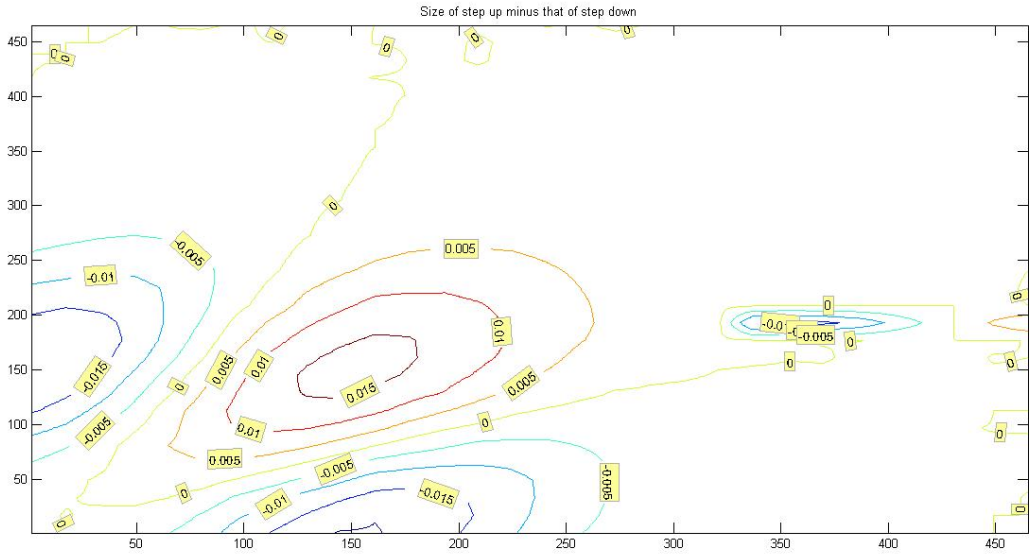

Figure 2.6: The difference of expected set sizes between set from the step-up method and the step-down method.

is so small (at most 0.1 in three-population case) that it would be hard to convince people to use it, considering so many more assumption they need.

Although I spent almost two years in this project and got no good result, it serves a very big lesson for me. It teaches me the correct way of doing research: every time we have a new method, first do some simulation to see if the improvement over the existing method is big enough to worth further attention. If I conducted the simulation at the beginning I would not have spent that two years trying to prove the relevant theorems and the write-ups. Luckily I stopped in time to work on the next project, which will be introduced in the next few chapters.

CHAPTER III

Set valued DTR

3.1 Introduction

In practice, clinicians offer treatment according to their patients' individual characteristics at each decision point. Clinicians collect and measure the information from the patient and recommend the category and dose of treatment correspondingly. Dynamic Treatment Regimes (DTRs) are sequences of decision rules that map from the patient's information to a recommended treatment. A DTR is analogous to a policy in reinforcement learning and to a controller in control theory. The ideas behind DTRs extend beyond medicine, appearing in education, marketing and economics, among other fields.

Most of the current research on DTRs focuses on constructing decision rules that take as input a patient's information and output one recommended treatment. However this is at odds with clinical decision support systems in which the clinician provides a set of treatments; this set excludes treatments known to be less effective given the patient's information. Examples include DXplain (*Barnett et al.*, 1987), Iliad (*Warner Jr*, 1989), RECONSIDER (*Blois et al.*, 1981) and QMR (*Miller and Masarie Jr*, 1989). Currently the set of treatments is informed by clinical expertise, biological theory, and—in the case where sequential decisions regarding treatment are required—only an indirect use of data (see (*Fortney et al.*, 2010) and (*Trivedi et al.*,

2014) for example).

At the beginning of this chapter, we consider the simple setting where there are just two decision points, and we assume that there are two (and just two) possible treatments at each stage. In the last section we will discuss the case when there are more than two treatment per stage This work borrows ideas from a variety of areas, including the area of *multiple comparisons with the best*. *Robins et al.* (2014) recommend that the first-stage treatment effect be a data-dependent parameter and work from the field of computer science that advocates assuming that the least effective in the set of treatments at the second stage may be selected. As will be both conceptually and mathematically explained in detail in section 3.2, this prospective has a bearing on how we define the effect of a non-final stage treatment: instead of assuming that patients will take the best corresponding future treatment, we assume that they will take the true worst corresponding future treatment in the recommended set. (The concept and construction of the recommended set will also be discussed in detail in section 3.2.) As we will see, the recommended set is random (i.e., depends on the data), and the parameters to be estimated are also random. Further discussion can be found in section 3.2.

A set valued DTR is different from a traditional DTR in two respects. First, in many cases we do not have enough evidence that the treatment recommended by the DTR is better than the remaining treatments. A set-valued DTR, instead of always recommending only one treatment at each decision point, will recommend a set of treatments. The treatments in this set are those we do not have enough statistical evidence on to distinguish from the best treatment. When one of the treatments is significantly better than the rest, the set will contain only this treatment.

The second difference is that we evaluate a non-final stage treatment in a different way. Traditional DTRs, using the dynamic programming idea, define the *effect* of a non-final stage treatment as the expected final outcome a patient would have, if he or

she took this treatment at the stage in question and took the best treatment at future stages. Our way of evaluating a treatment differs because we cannot assume that the patient will actually take the best treatment at all future stages. This assumption becomes even more unrealistic if the recommended set of treatments at some future stage contains several treatments. Thus, as will be discussed in section 3.2, we will define the effect of a treatment to be the expected final outcome if the patient takes this treatment at the current stage, and takes the worst from the set of recommended treatments at each future stages. The goal of recommending a set of treatments is therefore not to select the "best" treatment but rather to screen out treatments that are believed to be bad.

According to *Laber et al.* (2014a), when we cannot distinguish between the effects of several treatments, we should provide all such treatments to the clinician. However, as detailed in the following section, because we are providing a set of treatments at each stage, we do not know which one the patient will choose in the future. Therefore, our definition of the effect of a non-final stage treatment is different from that in *Laber et al.* (2010).

Other studies approach treatment regimes differently. *Fard and Pineau* (2009) discuss the construction of sets of Markov Decision Processes (MDPs). Instead of trying to identify the best policy (i.e., the function mapping a state to an action), they form a set of polices. Each policy in the set has expected discounted reward no less than $1-\epsilon$ times the largest expected discounted reward among all policies, where ϵ is a user-specified value.

Laber et al. (2014a) discuss set-valued DTRs from yet another point of view. In their work, they assume that there are two outcomes of interest. Instead of considering a linear combination of them as one single outcome, they construct the decision rule at each stage so as to provide a set of recommended treatments. A treatment is chosen for the recommended set if each of its interested outcomes is either better or not too

much worse than for the other treatments. This procedure can deal with the case where the two interested outcome have different units (e.g., the amount of reduction in blood pressure on the one hand and in blood fat on the other). Note, however, that giving a precise meaning to "not too much worse" requires a user specified threshold.

The main contribution of this work is twofold. First, we for the first time consider set valued DTRs when the interested outcome is a scaler without any user-specified threshold. To accommodate the set-valued DTRs we give a new definition of the effect of a non-final stage treatment, which usually is a random quantity instead of a fixed quantity. Second, we propose a new scientific goal of DTRs: instead of a DTR attempting to pin down the true best treatment, it should screen out significantly worse treatments for a given patient.

The rest of the chapter is organized as follows: in section 3.2 we introduce set valued DTRs and argue the coherent issue. In section 3.3 we provide a method for constructing the confidence intervals of the desired estimand and the corresponding asymptotic properties. In section 3.4 we examine the finite sample performance of the proposed confidence interval using simulated data. In section 3.5 we apply our method to data from an ADHD study. Section 3.6 provides a discussion of the open problems and future work. In the last section we will discuss the procedure of constructing the recommended set when there are more than two treatments at each stage. In the appendix we introduce traditional DTRs in detail and provide related proofs as well as more simulations.

3.2 Set-valued DTRs and the construction of the recommended sets

We suppose that at each stage there are two available treatments, denoted as $A_t \in \{-1, 1\}$ for $t = 1, 2$. We assume the data is from a sequentially randomized trial

(**?**), also known as a sequential multiple assignment randomized trial (*Murphy*, 2005a). For each patient, we observe a sequence (X_1, A_1, X_2, A_2, Y). Here X_t $(t = 1, 2)$ is the patient's covariate before the assignment of treatment A_t at stage t. The primary outcome, Y, is a scalar, coded such that a higher value indicates a better result. We let H_t denote the history before assigning treatment A_t, thus $H_1 = X_1$ and $H_2 = (X_1, A_1, X_2)$. At time t the clinician makes a decision based on H_t.

A traditional DTR in the two stage setting is composed of a pair of decision rules, $\boldsymbol{\pi}_{tra} = (\pi_{1,tra}, \pi_{2,tra})$, where $\pi_{t,tra}$ is a map from the domain of H_t to the domain of A_t, i.e., at each decision time point t, the output of $\pi_{t,tra}$ is a single treatment from the domain of A_t.

A set-valued DTR differs from a traditional DTR in that the range of a set-valued DTR is different. Specifically, a set-valued DTR is a pair of functions, $\boldsymbol{\pi}_{set} = (\pi_{1,set}, \pi_{2,set})$, where $\pi_{t,set}$ is a map from the domain of H_t to $Power(A_t)$, which is the power set of the domain of A_t but with $\emptyset$ excluded. For example, if A_t can take values in $\{-1, 1\}$, then $Power(A_t) = \{\{1\}, \{-1\}, \{1, -1\}\}$.

An optimal traditional DTR π_{tra}^{opt} is such that $\mathbb{E}^{\pi_{tra}^{opt}} Y = \sup_{\pi_{tra}} \mathbb{E}^{\pi_{tra}} Y$. It is a sequence of decision rules, which if implemented with the population of patients, will have the largest expected outcome. In contrast, *optimal* in the contest of set-valued DTRs may be defined in a variety of ways. We use a "worst-case" definition. Our worst-case definition allows for the fact, that given a set of treatments to choose between, the clinician/patient may select the worst-performing treatment in that set. In order to explain this definition, first we need to analyze how we compare treatment effects.

In order to compare treatments, we need a criterion: at each stage we compare treatments based on the Q-function for that stage. The Q-function for stage 2 is just

the conditional expectation

$$Q_2(h_2, a_2) \quad := \quad \mathbb{E}(Y|H_2 = h_2, A_2 = a_2) \tag{3.1}$$

Thus, $Q_2(h_2, a_2)$ is the expected outcome for patients with $H_2 = h_2$ who are provided $A_2 = a_2$ at the second stage. The parametric form of Q_2 can be defined as

$$Q_2(H_2, A_2; \theta_2) := H_{2,0}^T \theta_{2,0} + H_{2,1}^T \theta_{2,1} A_2 \tag{3.2}$$

For each h_2 we use an estimator of $Q_2(h_2, a_2)$ and ideas from literature on Multiple Comparisons with the Best (*Hsu*, 1996) to construct a recommended set $\hat{S}_2(h_2)$ of second-stage treatments. Each set contains the stage 2 treatments that can not be differentiated from the best stage 2 treatment.

Before introducing the construction of $\hat{S}_2(h_2)$, we define the second-stage estimand

$$\theta_2^* \quad := \quad \arg\min_{\theta_2} P(Y - Q_2(H_2, A_2; \theta_2))^2 \tag{3.3}$$

Here θ_2^* can be interpreted as the true second-stage regression coefficient. At stage 2 for a patient with $H_2 = h_2$, we are interested in $h_{2,1}^T \theta_{2,1}^*$, which is half the difference between the effect of treatment $A_2 = 1$ and the effect of treatment $A_2 = -1$. The estimator of θ_2^* is defined as $\hat{\theta}_2 := \arg\min_{\theta_2} \mathbb{P}_n(Y_2 - Q_2(H_2, A_2; \theta_2))^2$. The limiting distribution of $\sqrt{n} h_{2,1}^T (\hat{\theta}_{2,1} - \theta_{2,1}^*)$ is always normal; a proof of this fact can be found in the online supplementary materials.

To form $\hat{S}_2(h_2)$, we proceed as follows. First we construct two one-sided confidence intervals for $h_{2,1}^T \theta_{2,1}^*$ of the form $(-\infty, u)$ and $(l, +\infty)$, respectively. If $u > 0$, meaning we cannot reject the hypothesis $h_{2,1}^T \theta_{2,1}^* > 0$, we include treatment 1 in $\hat{S}_2(h_2)$. Similarly if $l < 0$, meaning we cannot reject $h_{2,1}^T \theta_{2,1}^* < 0$, we include treatment -1 in $\hat{S}_2(h_2)$.

Note that another way of constructing $\hat{S}_2(h_2)$ is to consider one two-sided confidence interval (say, $(\tilde{l}, \tilde{u})$) for $h_{2,1}^T \theta_{2,1}^*$ instead of two one-sided confidence intervals. In this variant, $\hat{S}_2(h_2)$ would contain both treatments if $0 \in (\tilde{l}, \tilde{u})$ and would contain only $sign(\tilde{l})$ otherwise (in this case $sign(\tilde{l}) = sign(\tilde{u})$).

Although the Lebesgue measure of the two one-sided confidence intervals is larger than that of the two-sided confidence interval (the former having infinite Lebesgue measure), the recommended set constructed from one-sided confidence intervals is less conservative than that constructed from two-sided confidence intervals, in the sense that it will contain the true inferior treatment with smaller probability. Indeed, although they both contain the true better treatment with probability no less than the given value (see (*Hsu*, 1996)), the probability of each treatment being included in the alternative $\hat{S}_2(h_2)$ constructed from a two-sided confidence interval will always be no smaller than the corresponding probability for the $\hat{S}_2(h_2)$ constructed from one-sided confidence intervals, since it is easy to show that we always have $\tilde{l} < l$ and $\tilde{u} > u$. In other words, the approach that uses one-sided confidence intervals to construct $\hat{S}_2(h_2)$ has greater power, and that is the approach we take in this chapter.

As the limiting distribution of $\sqrt{n}(\hat{\theta}_2 - \theta_2^*)$ is always normal, we can use a one sided t-test. Specifically, we consider the asymptotic pivot

$$\sqrt{n}(h_{2,1}^T \hat{\theta}_{2,1} - h_{2,1}^T \theta_{2,1}^*) / \sqrt{h_{2,1}^T \hat{\Sigma}_{21,21} h_{2,1}} \tag{3.4}$$

where $\hat{\Sigma}_{21,21}$ is the submatrix of $(\mathbb{P}_n B_2 B_2^T)^{-1} \mathbb{P}_n B_2 B_2^T (Y - B_2^T \hat{\theta}_{2,1})^2 (\mathbb{P}_n B_2 B_2^T)^{-1}$ corresponding to the plug-in estimator of the asymptotic variance of $\mathbb{V}_n := \sqrt{n}(\hat{\theta}_{2,1} - \theta_{2,1}^*)$; thus $\sqrt{h_{2,1}^T \hat{\Sigma}_{21,21} h_{2,1}}$ is the estimated standard deviation of $\sqrt{n}(h_{2,1}^T \hat{\theta}_{2,1} - h_{2,1}^T \theta_{2,1}^*)$. Note that (3.4) asymptotically follows the standard normal distribution. Let $z_{1-\alpha}$ be the $1-\alpha$ quantile of a standard normal distribution. Then we have two $(1-\alpha)$-level one-sided confidence interval for $h_{2,1}^T \theta_{2,1}^*$: $(-\infty, h_{2,1}^T \hat{\theta}_{2,1} + z_{1-\alpha} \sqrt{h_{2,1}^T \hat{\Sigma}_{21,21} h_{2,1}} / \sqrt{n})$

and $(h_{2,1}^T\hat{\theta}_{2,1} - z_{1-\alpha}\sqrt{h_{2,1}^T\hat{\Sigma}_{21,21}h_{2,1}}/\sqrt{n}, +\infty)$. We construct $\hat{S}_2$ as follows: we include 1 in $\hat{S}_2$ if $h_{2,1}^T\hat{\theta}_{2,1} > -z_{1-\alpha}\sqrt{h_{2,1}^T\hat{\Sigma}_{21,21}h_{2,1}}/\sqrt{n}$, and we include -1 in $\hat{S}_2$ if $h_{2,1}^T\hat{\theta}_{2,1} < z_{1-\alpha}\sqrt{h_{2,1}^T\hat{\Sigma}_{21,21}h_{2,1}}/\sqrt{n}$.

Note that $\hat{S}_2(h_2)$ contains both treatments if and only if

$$\sqrt{n}|h_{2,1}^T\hat{\theta}_{2,1}|/\sqrt{h_{2,1}^T\hat{\Sigma}_{21,21}h_{2,1}} \leq z_{1-\alpha}.$$

Alternatively if we define $T(h_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) = n(h_{2,1}^T\hat{\theta}_{2,1})^2/h_{2,1}^T\hat{\Sigma}_{21,21}h_{2,1}$ and $\chi = \chi(\alpha) = z_{1-\alpha}^2$, we see that $\hat{S}_2(h_2)$ contains only one treatment if and only if $T(h_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) > \chi$; it contains both treatments if and only if $T \leq \chi$. This is useful for the decomposition of $\sqrt{n}c^T(\hat{\theta}_1 - \hat{\theta}_1^*)$.

Next we consider stage 1. To define the Q-function for this stage, let $s_2(\cdot)$ be a deterministic function that maps the domain of H_2 to $Power(A_2)$. Define

$$Q_1(h_1, a_1; s_2) \;:=\; \mathbb{E}\left[\min_{a_2 \in s_2(H_2)} Q_2(H_2, a_2)\Big| H_1 = h_1, A_1 = a_1\right] \tag{3.5}$$

Its parametric form can be defined by

$$Q_1(H_1, A_1; \theta_1(s_2), s_2) \;:=\; H_{1,0}^T\theta_{1,0}(s_2) + H_{1,1}^T\theta_{1,1}(s_2)A_1 \tag{3.6}$$

where $\theta_1(s_2) := (\theta_{1,0}(s_2)^T, \theta_{1,1}(s_2)^T)^T$.

We also define the predicted outcome

$$\hat{Y}^*(s_2) \;:=\; \min_{a_2 \in s_2} Q_2(H_2, a_2; \theta_2^*) \tag{3.7}$$

Finally, for any second stage treatment regime $\hat{S}_2$, we define our estimand

$$\hat{\theta}_1^* = \hat{\theta}_1^*(\hat{S}_2) \;:=\; \left\{\arg\min_{\theta_1(s_2)} P(\hat{Y}(s_2) - Q_1(H_1, A_1; \theta_1(s_2), s_2))^2\right\}\Big|_{s_2 = \hat{S}_2} \tag{3.8}$$

where $\theta_1^*(\hat{S}_2) := (\theta_{1,0}^*(\hat{S}_2)^T, \theta_{1,1}^*(\hat{S}_2)^T)^T$.

Remark III.1. In (3.7), $\hat{Y}^*$ can be interpreted as the expected true final outcome for a patient if he or she had taken the true worse treatment in s_2 at the second stage.

Remark III.2. We can see in (3.8) that the estimand, $\hat{\theta}_1^*(\hat{S}_2)$ depends on $\hat{S}_2$, which is a random quantity. This follows the idea of *Robins et al.* (2014). We need to consider the case where patients choose the true worse treatment in the recommended set (if the set contains more than one treatment), so we need to consider the randomness of the recommended set. For notational convenience, we will denote $\hat{\theta}_1^*(\hat{S}_2)$ by $\hat{\theta}_1^*$.

Remark III.3. We can interpret $\hat{\theta}_1^*$ as the true first-stage regression coefficient if the second-stage recommended set is $\hat{S}_2$. At stage 1 for a patient with $H_1 = h_1$, under the assumption that the second-stage recommended set is $\hat{S}_2$, the patient's expected final outcome is $h_{1,0}^T\hat{\theta}_{1,0}^* + h_{1.1}^T\hat{\theta}_{1,1}^* A_1$. Thus we are interested in $h_{1,1}^T\hat{\theta}_{1,1}^*$, which is half the difference between the effects of the treatments $A_1 = 1$ and $A_1 = -1$.

Remark III.4. In stage 1 we construct two one-sided confidence intervals for $h_{1,1}^T\hat{\theta}_{1,1}^*$ similarly to how we constructed, in stage 2, the confidence intervals for $h_{2,1}^T\theta_{2,1}^*$. The estimator is defined as

$$\hat{\theta}_1 := \arg\min_{\theta_1} \mathbb{P}_n(\tilde{Y}_s - (H_{1,0}^T\theta_{1,0} + H_{1,1}^T\theta_{1,1}A_1))^2$$

where $\tilde{Y}_s = \min_{a_2 \in \hat{S}_2(H_2)} Q_2(H_2, a_2; \hat{\theta}_2)$, which is the estimator of $\hat{Y}^* = \min_{a_2 \in \hat{S}_2(H_2)} Q_2(H_2, a_2; \theta_2^*)$. The difficulty is that $\sqrt{n}h_{1,1}^T(\hat{\theta}_{1,1} - \hat{\theta}_{1,1}^*)$ might not always converge to a normal distribution and might have different limiting distribution from its bootstrap analog $\sqrt{n}h_{1,1}^T(\hat{\theta}_{1,1}^{(b)} - \hat{\theta}_{1,1})$. Thus our confidence interval for $\sqrt{n}h_{1,1}^T(\hat{\theta}_{1,1} - \hat{\theta}_{1,1}^*)$ from quantiles of $\sqrt{n}h_{1,1}^T(\hat{\theta}_{1,1}^{(b)} - \hat{\theta}_{1,1})$ will have poor coverage in some scenario. (Details of how this naive bootstrap method is carried out can be found at the beginning of section 3.4, the simulation study).

An example of the poor coverage of the naive bootstrap approach can be seen in

the simulation model discussed in *Laber et al.* (2010), Working under a generative model, we can show that if the significance level is set at 95%—i.e., we want to include the true better first-stage treatment in the first-stage recommended set with probability no less than 95%—then the recommended set constructed from the naive bootstrap method will include the true better treatment with probability around only 87%. To deal with this problem, we use the notion of Adaptive Confidence Interval (ACI), introduced by *Laber et al.* (2010), to construct a confidence interval for $\sqrt{n}h_{1,1}^T(\hat{\theta}_{1,1} - \hat{\theta}_{1,1}^*)$. The downside to this procedure, as will be shown in section 3.5, is that the recommended set constructed by our method is conservative, in the sense that it sometimes contains the true better treatment with probability larger than the specified value. The procedure and properties of ACI will be provided in the next section.

To summarize, our algorithm is as follows:

1. Calculate the second-stage regression: $\hat{\theta}_2 := \arg\min_{\theta_2} \mathbb{P}_n(Y_2 - Q_2(H_2, A_2; \theta_2))^2$.

1.1 Construct the second-stage confidence set: $\hat{S}_2(H_2) = \{-1, 1\}$ if we can reject neither that $A_2 = 1$ is better than $A_2 = -1$ nor that $A_2 = -1$ is better than $A_2 = 1$, and $\hat{S}_2(H_2) = \{sign(H_{2,1}^T\hat{\theta}_2)\}$ otherwise.

2. Predict the second-stage outcome: $\tilde{Y}_s = Y_1 + \min_{a_2 \in \hat{S}_2(H_2)} Q_2(H_2, a_2; \hat{\theta}_2)$.

3. Calculate the first-stage regression: $\hat{\theta}_1 := \arg\min_{\theta_1} \mathbb{P}_n(\tilde{Y}_s - Q_1(H_1, A_1; \theta_1))^2$.

3.1 Constructing stage 1 confidence set: $\hat{S}_1(H_1) = \{-1, 1\}$ if we can reject neither that $A_1 = 1$ is better than $A_1 = -1$ nor that $A_1 = -1$ is better than $A_1 = 1$, and $\hat{S}_1(H_1) = \{sign(H_{1,1}^T\hat{\theta}_1)\}$ otherwise.

3.3 Role of the ACI in the construction of the recommended set

In this section we focus on the ACI in order to describe in greater detail how to construct the recommended set of treatments. We continue to consider the case where there are two stages and, at each stage, there are two available treatments, denoted as -1 and 1.

Recall that at stage 2, for a patient with $H_2 = h_2 = (h_{2,0}^T, h_{2,1}^T)^T$, the true effect of treatment $A_2 = 1$ is $h_{2,0}^T\theta_{2,0}^* + h_{2,1}^T\theta_{2,1}^*$, while the true effect of treatment $A_2 = -1$ is $h_{2,0}^T\theta_{2,0}^* - h_{2,1}^T\theta_{2,1}^*$. We are interested in the quantity $h_{2,1}^T\theta_{2,1}^*$, which is half of the difference between effects of treatments 1 and -1. If we define $c = (\mathbf{0}^T, h_{2,1}^T)^T$, where $\mathbf{0}^T$ is a zero vector with the same length as $h_{2,0}^T$, then we can write $h_{2,1}^T\theta_{2,1}^* = c^T\theta_2^*$. If we cannot reject $c^T\theta_2^* \geq 0$, we include 1 in the recommended set; if we cannot reject $c^T\theta_2^* \leq 0$, we include -1 in the recommended set. Details of this hypothesis testing were introduced in section 3.2.

All the difficulty lies in the construction of the first-stage confidence interval. For a patient with $H_1 = h_1$, we would like to construct two one-sided confidence intervals for $h_{1,1}^T\hat{\theta}_{1,1}^* = c^T\hat{\theta}_1^*$, which is half the difference between the effects of $A_1 = 1$ and $A_1 = -1$, where $c = (\mathbf{0}^T, h_{1,1}^T)^T$. We want to construct confidence intervals based on $c^T\hat{\theta}_1$. however, as we will see below, $\sqrt{n}(\hat{\theta}_1^{(b)} - \hat{\theta}_1)$ and $\sqrt{n}(\hat{\theta}_1 - \hat{\theta}_1^*)$ might have different limiting distributions. Therefore, this naive bootstrap method will perform poorly, as the simulation study in section 3.4 will show.

For $t = 1, 2$, we define $B_t := (H_{t,0}^T, A_t H_{t,1}^T)^T, \Sigma_{t,\infty} := PB_tB_t^T$ and $\hat{\Sigma}_t := \mathbb{P}_n B_t B_t^T$. We assume $\hat{\Sigma}_t$ is invertible. Using standard methods it can be shown that $\mathbb{V}_n := \sqrt{n}(\hat{\theta}_2 - \theta_2^*)$ is asymptotically normal with mean zero and variance-covariance matrix

$$\mathbf{\Omega} = (PB_2B_2^T)^{-1}PB_2B_2^T(Y - B_2^T\theta_2^*)^2(PB_2B_2^T)^{-1}.$$

Let $\Sigma_{21,21}$ denote the submatrix of $\mathbf{\Omega}$ corresponding the limiting asymptotic covariance of $\sqrt{n}(\hat{\theta}_{2,1} - \theta^*_{2,1})$ and let $\hat{\Sigma}_{21,21}$ be the corresponding plug-in estimator. Then $\hat{\theta}_1 = \hat{\Sigma}_1^{-1}\mathbb{P}_n B_1 \tilde{Y}_s$. We can decompose $\sqrt{n}(\hat{\theta}_1 - \hat{\theta}^*_1)$ as

$$\mathbb{S}_n + \hat{\Sigma}_1^{-1}\mathbb{P}_n B_1(\mathbb{U}_n + \mathbb{O}_n) \tag{3.9}$$

where

$$\begin{aligned}
\mathbb{S}_n &= \hat{\Sigma}_1^{-1}\sqrt{n}\mathbb{P}_n B_1[Y_1 + H_{2,0}^T\theta^*_{2,0} - |H_{2,1}^T\theta^*_{2,1}|\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) \leq \chi) \\
&\qquad + H_{2,1}^T\theta^*_{2,1} sign(H_{2,1}^T\hat{\theta}_{2,1})\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) > \chi) \\
&\qquad - B_1^T\hat{\theta}^*_1 + H_{2,0}^T(\hat{\theta}_{2,0} - \theta^*_{2,0})] \qquad (3.10) \\
\mathbb{U}_n &= \sqrt{n}(|H_{2,1}^T\hat{\theta}_{2,1}| - |H_{2,1}^T\theta^*_{2,1}|)(2\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) > \chi) - 1) \qquad (3.11) \\
\mathbb{O}_n &= \sqrt{n}(|H_{2,1}^T\theta^*_{2,1}| - H_{2,1}^T\theta^*_{2,1} sign(H_{2,1}^T\hat{\theta}_{2,1}))\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) > \chi) \qquad (3.12)
\end{aligned}$$

with $\mathbb{1}(X)$ equals to 1 if X holds and 0 otherwise.

The term $\mathbb{S}_n$ is asymptotically normal, but $\mathbb{U}_n$ is non-smooth in $\hat{\theta}_{2,1}$. Therefore, following the idea of *Laber et al.* (2010), we divide the data into two parts: those for which we can not distinguish the two second-stage effects, and those for which we can. For the second group, we just perform the usual bootstrap, while for the first group, we let $\theta^*_{2,1}$ vary in its domain and take the supremum (or infimum) to be the upper (resp. lower) bound.

The term $\mathbb{O}_n$ is asymptotically zero but only if P is fixed; under some local alternative P_n it might not converge to zero. So we do the similar trick to $\mathbb{O}_n$ as well.

To be specific, the upper bound on $c^T\sqrt{n}(\hat{\theta}_1 - \hat{\theta}_1^*)$ is given by

$$
\begin{aligned}
& \mathcal{U}(c) \\
= \; & c^T\mathbb{S}_n + c^T\hat{\Sigma}_1^{-1}\mathbb{P}_n B_1\sqrt{n}(|H_{2,1}^T\hat{\theta}_{2,1}| - |H_{2,1}^T\theta_{2,1}^*|)\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) > \lambda_n) \\
& + c^T\hat{\Sigma}_1^{-1}\mathbb{P}_n B_1\sqrt{n}(|H_{2,1}^T\theta_{2,1}^*| - H_{2,1}^T\theta_{2,1}^* sign(H_{2,1}^T\hat{\theta}_{2,1}))\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) > \lambda_n) \\
& + \sup_{\gamma\in\mathbb{R}^{dim(\theta_{2,1}^*)}} \Big\{ c^T\hat{\Sigma}_1^{-1}\mathbb{P}_n B_1\Big[(|H_{2,1}^T(\mathbb{V}_n + \gamma) - |H_{2,1}^T\gamma|) \times \\
& (2\mathbb{1}(T(H_{2,1}, \mathbb{V}_n + \gamma, \hat{\Sigma}_{21,21}) > \chi) - 1) \\
& + (|H_{2,1}^T\gamma| - H_{2,1}^T\gamma sign(H_{2,1}^T(\mathbb{V}_n + \gamma))\mathbb{1}(T(H_{2,1}, \mathbb{V}_n + \gamma, \hat{\Sigma}_{21,21}) > \chi)\Big] \\
& \times\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) \le \lambda_n)\Big\}
\end{aligned} \tag{3.13}
$$

where λ_n is a sequence of deterministic numbers satisfying $\lim_{n\to\infty}\lambda_n = +\infty$ and $\lim_{n\to\infty}\lambda_n/n = 0$. If the sample size n is so small that $\lambda_n < \chi$, we force $\lambda_n = \chi$.

The lower bound $\mathcal{L}(c)$ is defined in a similar way, with "sup" replaced by "inf".

To see where $\mathcal{U}(c)$ comes from, notice that $\hat{\Sigma}_1^{-1}\mathbb{P}_n B_1(\mathbb{U}_n + \mathbb{O}_n)$, the second term in (3.9), is equal to

$$
\begin{aligned}
& \hat{\Sigma}_1^{-1}\mathbb{P}_n B_1\Big[\sqrt{n}(|H_{2,1}^T\hat{\theta}_{2,1}| - |H_{2,1}^T\theta_{2,1}^*|)(2\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) > \chi) - 1) \\
& + \sqrt{n}(|H_{2,1}^T\theta_{2,1}^*| - H_{2,1}^T\theta_{2,1}^* sign(H_{2,1}^T\hat{\theta}_{2,1}))\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) > \chi)\Big] \\
& \times\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) > \lambda_n) + \\
& \hat{\Sigma}_1^{-1}\mathbb{P}_n B_1\Big[\sqrt{n}(|H_{2,1}^T\hat{\theta}_{2,1}| - |H_{2,1}^T\theta_{2,1}^*|)(2\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) > \chi) - 1) \\
& + \sqrt{n}(|H_{2,1}^T\theta_{2,1}^*| - H_{2,1}^T\theta_{2,1}^* sign(H_{2,1}^T\hat{\theta}_{2,1}))\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) > \chi)\Big] \\
& \times\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) \le \lambda_n)
\end{aligned}
$$

In the first of the two terms, if $\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) > \lambda_n) = 1$ we must have that $\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) > \chi) = 1$ since $\lambda_n \ge \chi$. Therefore, the above term can

be simplified to

$$
\begin{aligned}
& \hat{\Sigma}_1^{-1}\mathbb{P}_n B_1\Big[\sqrt{n}(|H_{2,1}^T\hat{\theta}_{2,1}| - |H_{2,1}^T\theta_{2,1}^*|) + \sqrt{n}(|H_{2,1}^T\theta_{2,1}^*| - H_{2,1}^T\theta_{2,1}^* sign(H_{2,1}^T\hat{\theta}_{2,1}))\Big] \\
& \times \mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) > \lambda_n) + \\
& \hat{\Sigma}_1^{-1}\mathbb{P}_n B_1\Big[\sqrt{n}(|H_{2,1}^T\hat{\theta}_{2,1}| - |H_{2,1}^T\theta_{2,1}^*|)(2\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) > \chi) - 1) \\
& + \sqrt{n}(|H_{2,1}^T\theta_{2,1}^*| - H_{2,1}^T\theta_{2,1}^* sign(H_{2,1}^T\hat{\theta}_{2,1}))\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) > \chi)\Big] \\
& \times \mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) \leq \lambda_n) \\
= \; & \hat{\Sigma}_1^{-1}\mathbb{P}_n B_1\Big[\sqrt{n}(|H_{2,1}^T\hat{\theta}_{2,1}| - |H_{2,1}^T\theta_{2,1}^*|) + \sqrt{n}(|H_{2,1}^T\theta_{2,1}^*| - H_{2,1}^T\theta_{2,1}^* sign(H_{2,1}^T\hat{\theta}_{2,1}))\Big] \\
& \times \mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) > \lambda_n) + \\
& \hat{\Sigma}_1^{-1}\mathbb{P}_n B_1\Big[(|H_{2,1}^T(\mathbb{V}_n + \sqrt{n}\theta_{2,1}^*)| - |H_{2,1}^T\sqrt{n}\theta_{2,1}^*|) \times \\
& (2\mathbb{1}(T(H_{2,1}, \mathbb{V}_n + \sqrt{n}\theta_{2,1}^*, \hat{\Sigma}_{21,21}) > \chi) - 1) \\
& + (|H_{2,1}^T\sqrt{n}\theta_{2,1}^*| - H_{2,1}^T\sqrt{n}\theta_{2,1}^* sign(H_{2,1}^T(\mathbb{V}_n + \sqrt{n}\theta_{2,1}^*))) \times \\
& \mathbb{1}(T(H_{2,1}, \mathbb{V}_n + \sqrt{n}\theta_{2,1}^*, \hat{\Sigma}_{21,21}) > \chi)\Big] \\
& \times \mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) \leq \lambda_n)
\end{aligned}
$$

Notice that in the second term we have re-expressed $\sqrt{n}H_{2,1}^T\hat{\theta}_{2,1}$ as the sum of $H_{2,1}^T\mathbb{V}_n = H_{2,1}^T\sqrt{n}(\hat{\theta}_{2,1} - \theta_{2,1}^*)$ and $H_{2,1}^T\sqrt{n}\theta_{2,1}^*$. The quantity $H_{2,1}^T\sqrt{n}\theta_{2,1}^*$ characterizes the degree of non-regularity of $H_{2,1}^T\sqrt{n}(\hat{\theta}_1 - \hat{\theta}_1^*)$. One way to make the second term insensitive to local perturbation of $\theta_{2,1}^*$ is to replace $\sqrt{n}\theta_{2,1}^*$ with γ and take the supremum over all $\gamma \in \mathbb{R}^{dim(\theta_{2,1}^*)}$. (After making this replacement and taking the supremum, we have the form of $\mathcal{U}(c)$ in (3.13).) Doing this would give a regular upper bound. The reason we do not apply this trick to the first term is that those $H_{2,1}$ with $\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) > \lambda_n) = 1$ are highly likely to be regular: non-regularity occurs only when $H_{2,1}^T\theta_{2,1}^*$ is close to zero, which is unlikely the case when we have $\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) > \lambda_n) = 1$, i.e., $H_{2,1}^T\hat{\theta}_{2,1}$ has very large absolute value.

We remark that we use the function T both for estimation of $\hat{\theta}_1^*$ (those with comparison between T and χ) and for separation of data (those with comparison between T and λ_n). For the former the second input of T is also allowed to vary as a function of γ since in this situation T is a part of $\hat{\theta}_{2,1}$ while for the latter we just use $\sqrt{n}\hat{\theta}_{2,1}$ as the second input of T .

Before we formulate our result, we define

$$\begin{aligned} g_2(B_2, Y; \theta_2^*) &:= B_2(Y - B_2^T\theta_2^*). \\ g_1(B_1, H_2; \theta_1^*, \theta_2^*, \hat{\theta}_2, \hat{\Sigma}_{21,21}) &:= B_1\Big[H_{2,0}^T\theta_{2,0}^* - |H_{2,1}^T\theta_{2,1}^*|\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) \leq \chi) \\ &\quad + H_{2,1}^T\theta_{2,1}^* sign(H_{2,1}^T\hat{\theta}_{2,1})|\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) > \chi) \\ &\quad - B_1^T\hat{\theta}_1^*\Big] \end{aligned}$$

Assume the following:

(A1) The histories H_2, features B_1 and outcomes Y, satisfy the moment inequalities $P||H_2||^2||B_1||^2 < \infty$ and $PY^2||B_2||^2 < \infty$.

(A2) The matrices $\Sigma_{t,\infty}$ and $Cov(g_1, g_2)$ are strictly positive definite.

(A3) For any $s \in \mathbb{R}^{dim(\theta_{2,1}^*)}$, there exists a sequence of local alternatives P_n such that

(i) the P_n converge to P in the sense that

$$\int [\sqrt{n}(\mathrm{d}P_n^{1/2} - \mathrm{d}P^{1/2}) - \frac{1}{2}v_s \mathrm{d}P^{1/2}]^2 \to 0$$

for some real-valued measurable function v_s,

(ii) if $\theta_{2,n}^* := \arg\min_\theta P_n(Y - Q_2(H_2, A_2; \theta)^2$, then $\theta_{2,1,n}^* := \theta_{2,1}^* + s/\sqrt{n} + o(1/\sqrt{n})$, and

(iii) $P_n||H_2||^2||B_1||^2$ and $P_nY_2^2||B_2||^2$ are bounded sequences.

(A4) $\lim_{n\to\infty}\lambda_n = +\infty$ and $\lim_{n\to\infty}\lambda_n/n = 0$.

Define

$$\begin{aligned}
\hat{Y}^*_{1,n} &:= Y_1 + H^T_{2,0}\theta^*_{2,0,n} - |H^T_{2,1}\theta^*_{2,1,n}|\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) \leq \chi) \\
&\quad + H^T_{2,1}\theta^*_{2,1,n} sign(H^T_{2,1}\hat{\theta}_{2,1})|\mathbb{1}(T(H_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \hat{\Sigma}_{21,21}) > \chi) \\
\hat{\theta}^*_{1,n} &:= \arg\min_{\theta} P_n(\hat{Y}^*_{1,n} - Q_1(H_1, A_1; \theta))^2. \qquad (3.14)
\end{aligned}$$

Because in our proof we are using indicator functions, which are discontinuous at certain points, we make the mild assumption that P is absolutely continuous with respect to the Lebesgue measure m, i.e., we assume $P(A) = 0$ for all A with $m(A) = 0$. We call this the *absolute continuity assumption*. This assumption will be useful for the proof.

Now we show the limiting behavior. We define $\Sigma_{21,21,\infty}$ to be the limit of $\hat{\Sigma}_{21,21}$.

Theorem III.5.

1.

$$\begin{aligned}
&c^T\sqrt{n}(\hat{\theta}_1 - \hat{\theta}^*_1) \xrightarrow{d} \\
&c^T\mathbb{S}_\infty + P(c^T\Sigma^{-1}_{1,\infty}B_1H^T_{2,1}\mathbb{V}_1\mathbb{1}_{H^T_{2,1}\theta^*_{2,1}>0}) - P(c^T\Sigma^{-1}_{1,\infty}B_1H^T_{2,1}\mathbb{V}_1\mathbb{1}_{H^T_{2,1}\theta^*_{2,1}<0}) \\
&+c^T\Sigma^{-1}_{1,\infty}PB_1|H^T_{2,1}\mathbb{V}_1|(2\mathbb{1}(T(H_{2,1}, \mathbb{V}_1, \Sigma_{21,21,\infty}) > \chi) - 1)\mathbb{1}_{H^T_{2,1}\theta^*_{2,1}=0}
\end{aligned}$$

2. If, for each n, the underlying generative distribution is P_n which satisfying (A3),

then

$$
\begin{aligned}
&c^T\sqrt{n}(\hat{\theta}_1 - \hat{\theta}^*_{1,n}) \xrightarrow{d} \\
&c^T\mathbb{S}_\infty + c^T\Sigma^{-1}_{1,\infty}P(B_1H^T_{2,1}\mathbb{V}_1\mathbb{1}_{H^T_{2,1}\theta^*_{2,1}>0}) - c^T\Sigma^{-1}_{1,\infty}P(B_1H^T_{2,1}\mathbb{V}_1\mathbb{1}_{H^T_{2,1}\theta^*_{2,1}<0}) \\
&+c^T\Sigma^{-1}_{1,\infty}PB_1\Big[|H^T_{2,1}(\mathbb{V}_1+s)| - |H^T_{2,1}s|\Big] \times \\
&(2\mathbb{1}(T(H_{2,1},\mathbb{V}_1+s,\Sigma_{21,21,\infty})>\chi)-1)\mathbb{1}_{H^T_{2,1}\theta^*_{2,1}=0} \\
&-c^T\Sigma^{-1}_{1,\infty}PB_1\Big[(H^T_{2,1}s\cdot sign(H^T_{2,1}(\mathbb{V}_1+s)) - |H^T_{2,1}s|)\times \\
&\mathbb{1}(T(H_{2,1},\mathbb{V}_1+s,\Sigma_{21,21,\infty})>\chi)\Big]\mathbb{1}_{H^T_{2,1}\theta^*_{2,1}=0}
\end{aligned}
$$

3. The limiting distribution of $\mathcal{U}(c)$ *under both* P *and* P_n *is equal to*

$$
\begin{aligned}
&\mathcal{U}(c) \xrightarrow{d} \\
&c^T\mathbb{S}_\infty + c^T\Sigma^{-1}_{1,\infty}P(B_1H_{2,1}\mathbb{V}_1\mathbb{1}_{H^T_{2,1}\theta^*_{2,1}>0}) - c^T\Sigma^{-1}_{1,\infty}P(B_1H_{2,1}\mathbb{V}_1\mathbb{1}_{H^T_{2,1}\theta^*_{2,1}<0}) + \\
&\sup_{\gamma\in\mathbb{R}^{dim(\theta^*_{2,1})}}\Big[c^T\Sigma^{-1}_{1,\infty}P(B_1(|H^T_{2,1}(\mathbb{V}_1+\gamma)| - |H^T_{2,1}\gamma|)) \\
&\times(2\mathbb{1}(T(H_{2,1},\mathbb{V}_1+\gamma,\Sigma_{21,21,\infty})>\chi)-1)\mathbb{1}_{H^T_{2,1}\theta^*_{2,1}=0} \\
&-c^T\Sigma^{-1}_{1,\infty}PB_1(H^T_{2,1}\gamma sign(H^T_{2,1}(\mathbb{V}_1+\gamma)) - |H^T_{2,1}\gamma|)\times \\
&\mathbb{1}(\chi<T(H_{2,1},\mathbb{V}_1+\gamma,\Sigma_{21,21,\infty})\leq\lambda_n)\mathbb{1}_{H^T_{2,1}\theta^*_{2,1}=0}\Big]
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbb{S}_\infty \;=\; &\mathbb{G}_\infty\Big[c^T\Sigma^{-1}_{1,\infty}B_1[Y_1 + H^T_{2,0}\theta^*_{2,0} - |H^T_{2,1}\theta^*_{2,1}|\mathbb{1}(T(H_{2,1},\sqrt{n}\hat{\theta}_{2,1},\hat{\Sigma}_{21,21})\leq\chi) \\
&+H^T_{2,1}\theta^*_{2,1}sign(H^T_{2,1}\hat{\theta}_{2,1})\mathbb{1}(T(H_{2,1},\sqrt{n}\hat{\theta}_{2,1},\hat{\Sigma}_{21,21})>\chi) - B^T_1\hat{\theta}^*_1\Big] + c^T\Sigma^{-1}_{1,\infty}\Sigma_{12}\mathbb{V}_0
\end{aligned}
$$

Here $\mathbb{G}_\infty$ *is a tight Gaussian process in* $l^\infty(\mathcal{F}_2)$ *with covariance function* $Cov(\mathbb{G}_\infty f_1, \mathbb{G}_\infty f_2) = P(f_1 - Pf_1)(f_2 - Pf_2)$*, and* $\mathbb{V}_\infty = (\mathbb{V}_0, \mathbb{V}_1)$ *is the limiting distribution of* $\sqrt{n}(\hat{\theta}_2 - \theta^*_2)$*. Further,* χ *is the square of* $z_{1-\alpha}$ *the* $1-\alpha$ *quantile of a standard normal distribution.*

The limiting distribution of $\mathcal{L}(c)$ is similar, with "sup" replaced by "inf".

Note that the limiting distributions of $c^T\sqrt{n}(\hat{\theta}_1 - \hat{\theta}_1^*)$, $\mathcal{U}(c)$, and $\mathcal{L}(c)$ are the same in the case where $H_{2,1}^T\theta_{2,1}^* \neq 0$ with probability one. That is, when there is a large treatment effect for almost all patients then the upper (or lower) bound is tight. However, when there is a non-null subset of patients for whom there is no treatment effect, then the limiting distribution of the upper bound is stochastically larger than that of the limiting distribution of $c^T\sqrt{n}(\hat{\theta}_1 - \hat{\theta}_1^*)$. Thus, the ACI adapts to the setting in which all patients experience a treatment effect.

To form confidence intervals for $c^T\sqrt{n}(\hat{\theta}_1 - \hat{\theta}_1^*)$, the bootstrap distributions of $\mathcal{U}(c)$ and $\mathcal{L}(c)$ are used. The next result concerns the consistency of these bootstrap distributions. Let $\hat{\mathbb{P}}_n^{(b)}$ denote the empirical measure, that is, $\hat{\mathbb{P}}_n^{(b)} := n^{-1}\sum_{i=1}^n M_{n,i}\delta_{\mathcal{T}_i}$ for

$$(M_{n,1}, \ldots, M_{n,n}) \sim \text{Multinomial}(n, (1/n, \ldots, 1/n)).$$

We use the superscript (b) to signify that a functional has been replaced by its bootstrap analog, so that if $\omega := f(\mathbb{P}_n)$ then $\boldsymbol{\omega}^{(b)} := f(\hat{\mathbb{P}}_n^{(b)})$. Denote the space of bounded Lipschitz-1 functions on $\mathbb{R}^2$ by $BL_1(\mathbb{R}^2)$. Furthermore, let $\mathbb{E}_M$ and P_M denote the expectation and probability with respect to the bootstrap weights. The proofs of the following results appear after the proof of the limit results for $c^T\sqrt{n}(\hat{\theta}_1 - \hat{\theta}_1^*)$ and $\mathcal{U}(c)$:

Theorem III.6. *Assume (A1)-(A2), and fix $c \in \mathbb{R}^{dim(\theta_1^*)}$. Then $\mathcal{U}(c)$ and $\mathcal{U}^{(b)}(c)$ converge to the same limiting distribution in probability. That is,*

$$\sup_{v \in BL_1(\mathbb{R}^2)} |\mathbb{E}_v\mathcal{U}(c) - \mathbb{E}_{Mv}\mathcal{U}^{(b)}(c)|$$

converges in probability to zero.

Corollary III.7. *Assume (A1) (A2), fix $c \in \mathbb{R}^{dim(\hat{\theta}_1^*)}$, and let $\hat{u}$ and $\hat{l}$ denote the*

$1-\alpha$ *lower and upper quantile of* $\mathcal{U}^{(b)}(c)$. *Then*

$$P_M(c^T\hat{\theta}_1 - \hat{u}/\sqrt{n} \leq c^T\hat{\theta}_1^*) \geq 1-\alpha+o_P(1)$$

$$P_M(c^T\hat{\theta}_1 - \hat{l}/\sqrt{n} \geq c^T\hat{\theta}_1^*) \geq 1-\alpha+o_P(1)$$

Furthermore, if $P(H_{2,1}^T\theta_{2,1}^* = 0) = 0$, *then the above inequality can be strengthened to an equality.*

3.4 Simulation study

In this section we examine the small sample performance of the ACI method. We will compare the result with that from the naive bootstrap method, called the centered percentile bootstrap.

The naive bootstrap method, instead of getting bootstrap samples of $\mathcal{U}(c)$ defined in (3.13) to obtain $\mathcal{U}(c)^{(b)}$ in the paragraph above theorem III.6, bootstraps $c^T\sqrt{n}(\hat{\theta}_1 - \hat{\theta}_1^*)$ directly to obtain $c^T\sqrt{n}(\hat{\theta}_1^{(b)} - \hat{\theta}_1)$. Specifically, naive bootstrap method uses $\hat{u}_{naive}$ and $\hat{l}_{naive}$, the $1-\alpha$ lower and upper quantiles of $c^T\sqrt{n}(\hat{\theta}_1^{(b)} - \hat{\theta}_1)$ to construct two one-sided confidence interval for $c^T\hat{\theta}_1^*$: $(c^T\hat{\theta}_1 - \hat{u}_{naive}/\sqrt{n}, +\infty)$ and $(-\infty, c^T\hat{\theta}_1 - \hat{l}_{naive}/\sqrt{n})$.

In our simulation for the ACI method, we let $\lambda_n = \log(n)$. (Recall that λ_n is the critical value to determine whether we can distinguish the better second-stage treatment. More details can be found in the construction of $\mathcal{U}(c)$ in section 3.3.)

We use the same set of generative models as in *Laber et al.* (2010). The setting can be described as follows:

- $X_t \in \{-1, 1\}, A_t \in \{-1, 1\} \quad \text{for } t \in \{1, 2\}$
- $P(A_t = 1) = P(A_t = -1) = 0.5 \quad \text{for } t \in \{1, 2\}$
- $X_1 \sim \text{Bernoulli}(0.5), X_2|X_1, A_1 \sim \text{Bernoulli}(\text{expit}(\delta_1 X_1 + \delta_2 A_1))$

- $Y = \gamma_1 + \gamma_2 X_1 + \gamma_3 A_1 + \gamma_4 X_1 A_1 + \gamma_5 A_2 + \gamma_6 X_2 A_2 + \gamma_7 A_1 A_2 + \epsilon, \epsilon \sim N(0,1)$

Here $\text{expit}(x) = \exp(x)/(1+\exp(x))$. This class is parametrized by nine values, $\gamma_1, \ldots, \gamma_7, \delta_1, \delta_2$. The analysis model uses feature vectors defined by

$$\begin{aligned} H_{2,0} &= (1, X_1, A_1, X_1 A_1, X_2)^T, \\ H_{2,1} &= (1, X_2, A_1)^T. \\ H_{1,0} = H_{1,1} &= (1, X_1)^T \end{aligned}$$

Recall that our models are

$$\begin{aligned} Q_2(H_2, A_2; \theta_2) &= H_{2,0}^T \theta_{2,0} + H_{2,1}^T \theta_{2,1} A_2 \\ Q_1(H_1, A_1; \theta_1) &= H_{1,0}^T \theta_{1,0} + H_{1,1}^T \theta_{1,1} A_1 \end{aligned}$$

For both simulations, we use $n = 300$ samples. The number of bootstrap iteration is taken to be $B = 1000$. The significance level is set at 0.05. i.e., we want to include the better first-stage treatment with probability at least 95%.

Recall that in Q-learning, non-regularity occurs when two or more second stage treatments have almost the same expected final outcome. In our model class above, this occurs if the model generates histories for which $\gamma_5 A_2 + \gamma_6 X_2 A_2 + \gamma_7 A_1 A_2 \approx 0$. By manipulating the values of γ_i and δ_i, we can control (i) the probability of generating a patient history such that $\gamma_5 A_2 + \gamma_6 X_2 A_2 + \gamma_7 A_1 A_2 = 0$, (ii) the standardized effect size $\phi = E(\gamma_5 A_2 + \gamma_6 X_2 A_2 + \gamma_7 A_1 A_2)/\sqrt{Var(\gamma_5 A_2 + \gamma_6 X_2 A_2 + \gamma_7 A_1 A_2)}$ and (iii) the difference between the effects of $A_1 = 1$ and $A_1 = -1$. Note that (iii) is usually not a fixed quantity: recall that our definition of the effect of a first-stage treatment depends on the second-stage recommended set, which itself depends on the samples. For each generative model, we list the simulated difference between the effects of $A_1 = 1$ and $A_1 = -1$. The descriptions of the models are listed in table 3.1.

Table 3.1: Description of the simulation models using ACI

Index	γ	δ	p	ϕ	$2H_{1,1}\hat{\theta}^*_{1,1}$
One	$(0,0,0,0,0,0,0)^T$	$(0.5,0.5)^T$	1	0/0	0
Two	$(0,0,0,0,0.01,0,0)^T$	$(0.5,0.5)^T$	0	∞	$[-0.02,0.02]$
Three	$(0,0,-0.5,0,0.5,0,0.5)^T$	$(0.5,0.5)^T$	1/2	1.0	0
Four	$(0,0,-0.5,0,0.5,0,0.49)^T$	$(0.5,0.5)^T$	0	1.02	$[-0.02,0]$
Five	$(0,0,-0.5,0,1,0.5,0.5)^T$	$(0,0)^T$	1/4	1.41	0
Six	$(0,0,-0.505,0,1,0.49,0.5)^T$	$(0,0)^T$	0	1.43	$[-0.01,0]$
Seven	$(0,0,-0.495,0,1,0.5,0.49)^T$	$(0,0)^T$	0	1.43	$[-0.01,0]$

We compare the empirical performance of the ACI with the centered percentile bootstrap, the naive bootstrap method discussed before. Our primary interest is the probability of including the true better first-stage treatment in the recommended set. If the two first-stage treatments have the same effect, we consider the minimum of the two probabilities of including each of them. Recall that $2H_{1,1}\hat{\theta}^*_{1,1}$ is the difference between the effect of $A_1 = 1$ and the effect of $A_1 = -1$.

Also, we want to compare the width of the confidence intervals from two methods. Since we are constructing one-sided confidence intervals, we must make clear what we mean by the *width* of a one-sided confidence interval: this is defined to be the distance between the upper (or lower) bound and the point estimator. Recall that we are constructing upper and lower $1-\alpha$ confidence intervals for half of the difference between the effects of $A_1 = 1$ and $A_1 = -1$, and the two one-sided confidence intervals have the forms $(-\infty, u)$ and $(l, +\infty)$. We use the subscripts of ACI and BCI to indicate the corresponding u and l from ACI method and centered bootstrap method. We denote the common point estimator of this difference by $\hat{\theta}$. Then we are interested in the comparison between $u_{ACI} - \hat{\theta}$ and $u_{BCI} - \hat{\theta}$, as well as that between $\hat{\theta} - l_{ACI}$ and $\hat{\theta} - l_{BCI}$. We are also interested in r the average ratio of the length of the confidence interval coming from the ACI method to the corresponding quantity for the centered bootstrap method: $r = \mathbb{E}(u_{ACI} - l_{ACI})/(u_{BCI} - l_{BCI})$. The results are listed in table 3.2. More simulations can be found in the supplementary materials.

Table 3.2: Simulation results of the ACI

Index	P_{ACI}	P_{BCI}	$l_{1,ACI}$	$l_{1,BCI}$	$l_{2,ACI}$	$l_{2,BCI}$	r
1	0.9760	0.9230∗	0.2385	0.1748	0.2366	0.1716	1.3719
2	0.9905	0.9455	0.2386	0.1746	0.2369	0.1723	1.3707
3	0.9460	0.8820∗	0.2190	0.1965	0.1895	0.1475	1.1906
4	0.9560	0.8880∗	0.2188	0.1955	0.1896	0.1476	1.1903
5	0.9400	0.9110∗	0.2019	0.1904	0.1892	0.1674	1.0930
6	0.9440	0.9120∗	0.2023	0.1885	0.1909	0.1667	1.0929
7	0.9440	0.9110∗	0.2028	0.1890	0.1914	0.1672	1.0927

P=Probability of choosing the true better first stage treatment.
∗: Significantly smaller than 0.95 at 0.05 level.
$l_{1,ACI}$=$E(u_{ACI} - \hat{\theta})$, $l_{2,ACI}$=$E(\hat{\theta} - l_{ACI})$
$l_{1,BCI}$=$E(u_{BCI} - \hat{\theta})$, $l_{2,BCI}$=$E(\hat{\theta} - l_{BCI})$
$r = \mathbb{E}(u_{ACI} - l_{ACI})/(u_{BCI} - l_{BCI})$

We can see that, in all seven settings, the first-stage recommended sets from ACI method satisfy the goal: the true better first stage treatment is included with probability no smaller than 95%. By contrast, the naive bootstrap method fails in several settings especially in setting 3. This is the case where for patients who start with $A_1 = 1$, one second-stage treatment has much better effect than the other, so no non-regularity occurs. For patients who start with $A_1 = -1$, the two second-stage treatments have the same effect, so the second-stage recommended set will include both treatments with high probability; therefore, the method tends to underestimate the final expected outcome Y that a patient take the worse of the two treatments (their effects are the same so they are both the worse one). Thus naive bootstrap method underestimates the effect of $A_1 = -1$ and consequently includes it with small probability.

3.5 Analysis of the ADHD study

In this section we will illustrate our method by applying it to the data from the ADHD study by Pelham et al. For more information on this study, please refer to

Nahum-Shani et al. (2012b).

Attention-Deficit Hyperactivity Disorder (ADHD) is a neurodevelopmental psychiatric disorder treated mostly through counseling, lifestyle choices, and medication. Here we consider two particular approaches to treating ADHD at the first stage: (i) offering low-intensity behavioral intervention, and (ii) giving low-dosage medication. Our primary goal is to test the effectiveness of each of the two approaches as a first-stage treatment.

The design of the study is as follows. At the beginning of the study, half of the children were randomly assigned (with probability 0.5) to low-intensity behavioral intervention (coded as $A_1 = 1$) or low-dosage medication (coded as $A_1 = -1$). Starting at the eighth week, each child's response to the first-stage intervention was evaluated monthly until the end of the school year. A monthly rating from the Impairment Rating Scale (IRS) was used, along with an individualized list of target behaviors (ITB), to determine whether a child was a responder. A child was defined to be a non-responder if his/her average performance on the ITB was less than 75% and he/she was rated by teachers as impaired on the IRS in at least one domain. If the child was classified as a responder, he/she would remain in the first stage treatment and continue the assigned treatment. If the child was a non-responder, he/she would enter the second stage. Each of these children was re-randomized (with probability 0.5) to either augmenting the first-stage treatment option with the other type of treatment (i.e., adding medication if started with behavior intervention, and adding behavior intervention if started with medication, coded as $A_2 = 1$), or continuing the same treatment but with increasing dose/intensity (coded as $A_2 = -1$). The graph of the structure of this study is shown in figure 3.1.

The first step in using Q-learning is to estimate a regression model for the second stage. In this study, only non-responders have second stage treatment options. Two tailoring variables are used here: one is the first-stage treatment A_1, and the other

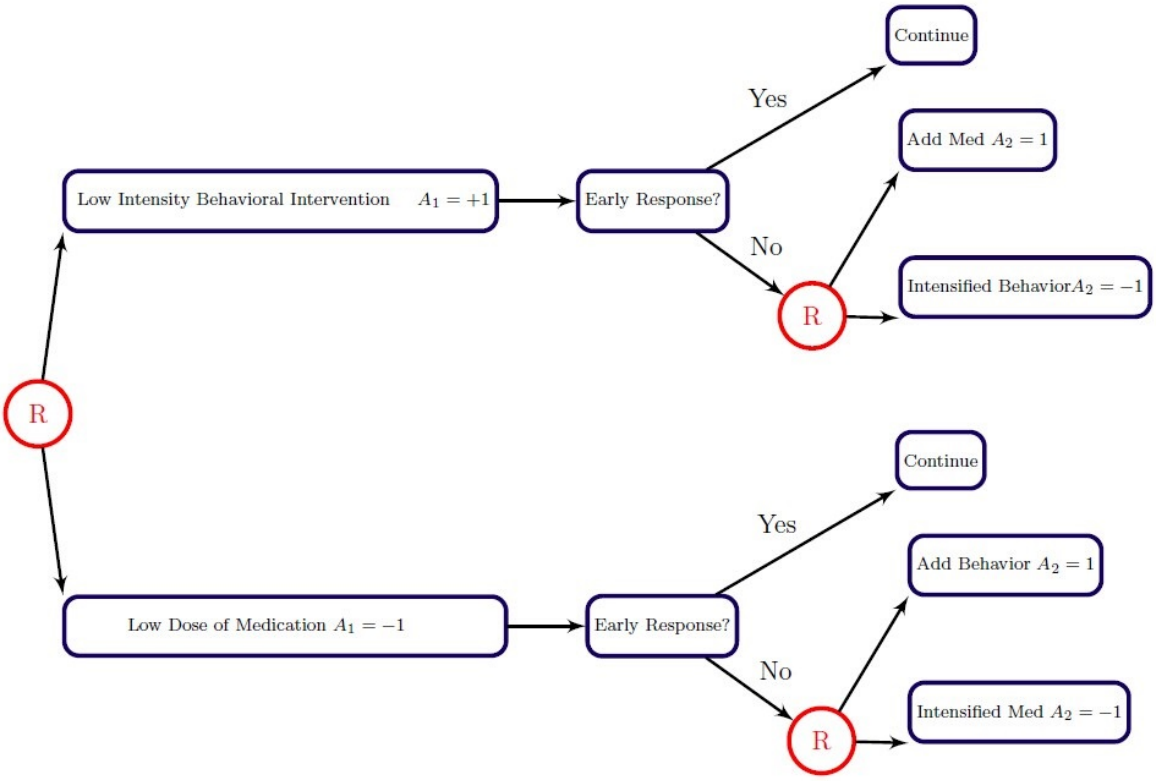

Figure 3.1: The design of the ADHD study.

Table 3.3: Descriptions of variables in ADHD

Covariate	Description
$X_{1,1}$	Medication prior to the first stage treatment. Binary: 1 for received and 0 for not received.
$X_{1,2}$	ADHD symptoms at the end of the previous school year. Continuous.
$X_{1,3}$	ODD. Binary: 1 for had ODD before and 0 otherwise.
$X_{2,1}$	The month during the school year at which the child showed inadequate response to the first stage treatment. Continuous.
$X_{2,2}$	Adherence. Binary: 1 for adherence and 0 otherwise.

ODD=Oppositional Defiant Disorder.

Table 3.4: Second-stage results for ADHD

A_1	$X_{2,2}$	CI_1	CI_2	$\hat{S}_2$
1	0	$(-\infty, -0.2282)$	$(-0.9783, +\infty)$	{-1}
1	1	$(-\infty, 0.7041)$	$(-0.0427, +\infty)$	{1,-1}
-1	0	$(-\infty, -0.1909)$	$(-1.2138, +\infty)$	{-1}
-1	1	$(-\infty, 0.6624)$	$(-0.1992, +\infty)$	{1,-1}

is the adherence to the first-stage treatment. The adherence is binary and can take either values 1 (for high adherence) or the value 0 (for low adherence). The meaning of low adherence depends on the type of treatment the child is started on: it means the treatment is received on fewer than 75% (respectively, 100%) of the days in the case of behavioral intervention (respectively, medication). The feature vectors we use for the second stage are $H_{2,0} = (1, X_{1,1}, X_{1,2}, X_{1,3}, A_1, A_1X_{1,1}, X_{2,1}, X_{2,2})^T$ and $H_{2,1} = (1, A_1, X_{2,2})^T$. The descriptions of the variables are given in table 3.3.

There are only four possible values for $H_{2,1}$. For each, we construct the upper and lower 95% one-sided confidence intervals of $H_{2,1}^T\theta_{2,1}^*$ (i.e., half of the difference between the effects of $A_2 = 1$ and $A_2 = -1$). The results are listed in table 3.4.

Recall that for each $h_{2,1}^T\theta_{2,1}^*$ we construct two one-sided confidence intervals with forms $(-\infty, u)$ and $(l, +\infty)$. We include $A_2 = 1$ in the second-stage recommended set $\hat{S}_2$ if $u \geq 0$ (i.e., we fail to reject the hypothesis that the difference between the effect of $A_2 = 1$ and that of $A_2 = -1$ is positive, meaning that we fail to reject that $A_2 = 1$ is better than $A_2 = -1$). Similarly we include $A_2 = -1$ in the second-stage recommended set $\hat{S}_2$ if $l \leq 0$.

We can see from table 3.4 that for children started with either of the first-stage treatments, if he/she adhered to it, we cannot distinguish the two second-stage treatments at the 95% confidence level; meanwhile for children started with either first stage treatment, if he/she didn't adhere to it an intensified first-stage treatment would be significantly better than adding the other first-stage treatment.

Our results are the same as those in *Nahum-Shani et al.* (2012b), in the sense of

Table 3.5: First-stage result for AHDH

$X_{1,1}$	CI_1	CI_2	BCI_1	BCI_2	$\hat{S}_1$
1	$(-\infty, 0.2090)$	$(-0.7275, +\infty)$	$(-\infty, 0.0662)$	$(-0.6477, +\infty)$	{1,-1}
0	$(-\infty, 0.3799)$	$(-0.2387, +\infty)$	$(-\infty, 0.3057)$	$(-0.1591, +\infty)$	{1,-1}

whether we can distinguish the two second-stage treatments for a given $H_{2,1}$. But the picture changes if we set the confidence level at 90% instead of 95%. In this case, for children who were started with behavioral intervention and adhered to it, our method shows that adding the other first-stage treatment is significantly better than intensification. By contrast, we found that the method used in *Nahum-Shani et al.* (2012b) is unable to distinguish these two second-stage options even at the 90% confidence level. The advantage in our method comes mainly from the use of one-sided confidence intervals instead of a two-sided confidence interval.

The second step in using the Q-learning is to construct the first-stage recommended set. At this stage we are interested in whether having a medication before entering the study will effect the recommended set. The feature vectors used here are $H_{1,0} = (1, X_{1,1}, X_{1,2}, X_{1,3})^T$ and $H_{1,1} = (1, X_{1,1})^T$. The results are listed in table 3.5.

In this table, CI means confidence interval from our method, while BCI means confidence interval from naive bootstrap method.

We can see that the confidence intervals from both methods inform no significant difference between the two first-stage treatments. This conclusion is the same as that in *Nahum-Shani et al.* (2012b). However, note that the "width" (i.e., the distance between the bound and the point estimator) of the confidence intervals from our method is larger than that from naive bootstrap method.

3.6 Conclusion and future work

In this chapter we discussed the non-regularity problem that occurs in the parameter estimation and one way of solving it. We demonstrated that the Adaptive Confidence Interval (ACI) method can handle the cases where non-regularities occur. However, as mentioned by *Laber et al.* (2010), the confidence interval of the first-stage-parameters estimation is conservative and therefore leads to the conservation of the recommended set. That is, when the effects of two first-stage treatments are the same, the probability of them being included in the recommended set is larger than the value we set. This is a natural result of the construction of the upper and lower bounds of the confidence intervals of the first-stage parameters. We can use a data dependent critical value for λ_n instead of always setting it to be $\log(n)$. We can also adopt the m-out-of-n bootstrap idea of *Chakraborty et al.* (2013) to decide the value of λ_n according to the data.

All sections in this chapter but the last one consider the scenario where there are two stages and two treatments at each stage. There are two directions in which to extend this scenario. One is to consider the case where there are three or more stages, possibly following the idea in *Laber et al.* (2010). The other direction is to consider the case where there are more than two available treatments at each stage. Suppose at stage t we have patient's history H_t and N_t available treatments, coded as $\{1, \ldots, N_t\}$. Our Q-function at each stage can be written as

$$Q_t(H_t, A_t; \theta_t) := \sum_{i=1}^{N_t} H_t^T \theta_{t,i} \mathbb{1}(A_t = i) \tag{3.15}$$

So for a patient with $H_t = h_t$, the treatment effects are $\theta = (\theta_1, \cdots, \theta_{N_t})$ where $\theta_i = h_t^T \theta_{t,i}$. Following the idea of *Hsu* (1996), for each treatment i, we are interested

in the $(N_t - 1)$-dimensional difference vector

$$\boldsymbol{\delta}^{(i)} = (\frac{\theta_i - \theta_1}{\sigma_{i,1}}, \cdots, \frac{\theta_i - \theta_{i-1}}{\sigma_{i,i-1}}, \frac{\theta_i - \theta_{i+1}}{\sigma_{i,i+1}}, \cdots, \frac{\theta_i - \theta_{N_t}}{\sigma_{i,N_t}}) \tag{3.16}$$

where $\sigma_{i,j}$ is the variance of $\hat{\theta}_i - \hat{\theta}_j$, the point estimator for the true difference in treatment effects between treatments i and j. The ith treatment is the best if and only if all the components of $\boldsymbol{\delta}^{(i)}$ are non-negative. So we need to form a simultaneous confidence interval for each component of $\boldsymbol{\delta}^{(i)}$.

In practice, we seldom know $\sigma_{i,j}$, and we have to use an estimator $\hat{\sigma}_{i,j}$ for it. The challenge is that at the first stage, the point estimator $\hat{\theta}_i$ is a non-smooth function of the data, so $\hat{\sigma}_{i,j}$ needs to be carefully designed.

3.7 Three treatment per stage case

3.7.1 Introduction

In practice, we often face the situation when there are more than two treatments at each stage. The reason and idea of constructing the recommended sets for each stage are the same as what have been discussed in the earlier sections. The key difference from the two-treatments-per-stage and hence the hard point, as will be discussed in detail in the following sub-sections, is that now we are comparing three treatments at a time as opposed to two. When there are only two treatments at hand, all we need to do is to construct the confidence interval for the difference of effects between these two treatments. Now for the situation where there are three, following the MCB idea we need to construct simultaneous confidence intervals for a two-dimensional vector, i.e., the differences of the effects between one of the three and the other two. Thus, how to "distribute" the "length ratio" of the two confidence intervals becomes a big issue. Again we will discuss it in detail when mentioning this issue.

3.7.2 Formulation of the problem

We start from the most simple model. Now we have two stages denoted as stage one and two. For stage one there are *three* available treatments, and for stage two, there are also three available treatments (for simplicity we assume no matter what you take at stage one, the number of available treatments in stage two are the same, i.e., three.)

Assume that we have patients' history H_t before we make the tth stage treatment recommendation and the three available treatments are denoted as 0,1 and 2. Below are our Q-learning model for stage two and its parametric form:

$$Q_2(H_2, A_2) := E(U|H_2 = h_2, A_2 = a_2) \tag{3.17}$$

$$Q_2(H_2, A_2; \beta_2) := H_{2,0}^T\beta_{2,0} + H_{2,1}^T\beta_{2,1}\mathbb{1}(A_2 = 1) + H_{2,2}^T\beta_{2,2}\mathbb{1}(A_2 = 2) \tag{3.18}$$

Remark:

1. We denote $\beta_2 = (\beta_{2,0}^T, \beta_{2,1}^T, \beta_{2,2}^T)^T$.

2. Another choice of parametrization for (3.18) is

$$Q_2(H_2, A_2; \beta_2) := H_{2,0}^T\beta_{2,0}\mathbb{1}(A_2 = 0) + H_{2,1}^T\beta_{2,1}\mathbb{1}(A_2 = 1) + H_{2,2}^T\beta_{2,2}\mathbb{1}(A_2 = 2) \tag{3.19}$$

which looks more "symmetric". But there is a potential problem with this form. In the form of (3.18), $H_{2,0}$ can include variables that do not interact with the treatments but in the form of (3.19) there is no such option. So throughout this section we will use (3.18).

3. In order to use the expression in (3.18), we need to put some constrain to make it "invariant" under different coding of the three treatments. To be specific, the general for of (3.18) is

$$Q_2(H_2, A_2; \beta_2) := H_{2,0}^T\beta_{2,0} + H_{2,1}^T\beta_{2,1}f_1(A_2) + H_{2,2}^T\beta_{2,2}f_2(A_2) \tag{3.20}$$

where the matrix

$$V_F = \begin{pmatrix} 1 & f_1(0) & f_2(0) \\ 1 & f_1(1) & f_2(1) \\ 1 & f_1(2) & f_2(2) \end{pmatrix}$$

has full rank. We need that for two investigators using two different choices of (f_1, f_2) will have the same conclusion (i.e., the estimator of $Q_2(H_2, A_2)$ should be the same from the two choices for all values of A_2, as well as the covariance matrix of this estimator). A necessary and and sufficient condition for this is given in the following lemma.

Lemma III.8. *Invariance to the coding of the treatments For any fixed dataset satisfying: $H_{2,1}$ is a sub-vector of $H_{2,0}$ and $H_{2,1} = H_{2,2}$, then for any $f = (f_1, f_2)$ satisfying that V_F defined above is of full rank, from (3.22) and (3.23) we get $\hat{\beta}_{2,f}$ and $\hat{\Sigma}_{2,f}$. Then neither*

$$H_{2,0}^T \hat{\beta}_{2,0} + H_{2,1}^T \hat{\beta}_{2,1} f_1(A_2) + H_{2,2}^T \hat{\beta}_{2,2} f_2(A_2)$$

nor

$$(H_{2,0}^T, H_{2,1}^T f_1(A_2), H_{2,2}^T f_2(A_2))^T \hat{\Sigma}_{2,f} (H_{2,0}^T, H_{2,1}^T f_1(A_2), H_{2,2}^T f_2(A_2))$$

depend on f, for all $A_2 = 0, 1, 2$.

The proof of this lemma can be found in the appendix. So throughout this section we will assume that $H_{2,1}$ is a sub-vector of $H_{2,0}$ and $H_{2,1} = H_{2,2}$. And we will use the coding of (3.18), which means we use $f_1(A_2) = \mathbb{1}(A_2 = 1)$ and $f_2(A_2) = \mathbb{1}(A_2 = 2)$.
4. Although we only consider the case of three treatments per stage, the scenario of more than three treatments can be dealt with similarly, although as we will see, the deduction will be more complicated.

We have two goals. First, for a patient at the second stage with $H_2 = h_2$, among three available treatments, we would like to provide a non-empty set of treatments

that is a subset of $\{0, 1, 2\}$, such that the true best second stage treatment will be included in this subset with probability no less than a pre-specified level, say, $1 - \alpha$. Our second goal is that for a patient at the first stage with $H_1 = h_1$, among three available treatments we would like to provide a non-empty set of treatments that is a subset of $\{0, 1, 2\}$ such that the true best first stage treatment will be included in this subset with probability no less than $1 - \alpha$. To compare treatments in order to define the "best" one we need to define the "effect" of a treatment. We will make this definition later.

The second stage estimand is defined as

$$\beta_2^* := \arg\min_{\beta_2} P(Y - Q_2(H_2, A_2; \beta_2))^2 \tag{3.21}$$

Here β_2^* can be interpreted as the true second-stage regression coefficient. At stage two we are interested in vector $(H_{2,1}^T\beta_{2,1}^*, H_{2,2}^T\beta_{2,2}^*)$. The first component is the true difference of effect between $A_2 = 1$ and $A_2 = 0$ while the second component is the true difference of effect between $A_2 = 2$ and $A_2 = 0$. If we have $H_{2,1}^T\beta_{2,1}^* \leq 0$ and $H_{2,2}^T\beta_{2,2}^* \leq 0$, then for this patient $A_2 = 0$ is the best. If we have $H_{2,1}^T\beta_{2,1}^* \geq 0$ and $H_{2,1}^T\beta_{2,1}^* - H_{2,2}^T\beta_{2,2}^* \geq 0$, then for this patient $A_2 = 1$ is the best. Finally if we have $H_{2,2}^T\beta_{2,2}^* \geq 0$ and $H_{2,1}^T\beta_{2,1}^* - H_{2,2}^T\beta_{2,2}^* \leq 0$, then for this patient $A_2 = 2$ is the best.

The estimator of β_2^* is

$$\hat{\beta}_2 := \arg\min_{\beta_2} \mathbb{P}_n(Y - Q_2(H_2, A_2; \beta_2))^2 \tag{3.22}$$

where $\mathbb{P}_n$ is the empirical measure. And the estimated variance matrix of $\hat{\beta}_2$ is

$$\big[(Y - B_2^T\hat{\beta}_2)^T(Y - B_2^T\hat{\beta}_2)/n\big](B_2B_2^T)^{-1} \tag{3.23}$$

where B_2^T is an n by n_2 second stage design matrix, with n being the number of observations and n_2 the dimension of H_2.

The procedure of constructing $\hat{S}_2$ is as follows. First we fix our significance level α. For each individual with $H_2 = h_2$, first we do the hypothesis test with the null being $A_2 = 0$ being the best, i.e., H_0 is $-h_{2,1}^T \beta_{2,1}^* \geq 0$ and $-h_{2,2}^T \beta_{2,2}^* \geq 0$, with significance level α (From now on every hypothesis test without specifying the significance level will have a default significance level α). If we fail to reject this hypothesis, we include $A_2 = 0$ in $\hat{S}_2$, otherwise we do not include it. The test statistic is $(-h_{2,1}^T \hat{\beta}_{2,1}, -h_{2,2}^T \hat{\beta}_{2,2})$. We fail to reject the null when both of these two elements are greater than some negative quantities. Details will be provided later. Similar procedures are performed for deciding whether to include $A_2 = 1$ and $A_2 = 2$.

Next we show the detail of testing whether $A_2 = 0$ is the best for patients with $H_2 = h_2$. Let $\hat{\sigma}_{2,1}^2(h_2) = h_{2,1}^T \hat{\Sigma}_{21,21} h_{2,1}$ be the estimated variance of $\sqrt{n} h_{2,1}^T \hat{\beta}_{2,1}$, where $\hat{\Sigma}_{21,21}$ is the estimated variance matrix of $\sqrt{n}\hat{\beta}_{2,1}$. Similarly define $\hat{\sigma}_{2,2}^2(h_2) = h_{2,2}^T \hat{\Sigma}_{22,22} h_{2,2}$. We also define $\hat{\sigma}_{2,12}(h_2) = h_{2,1}^T \hat{\Sigma}_{21,22} h_{2,2}$ where $\hat{\Sigma}_{21,22}$ is the estimated covariance matrix of $\sqrt{n} h_{2,1}^T \hat{\beta}_{2,1}$ and $\sqrt{n} h_{2,2}^T \hat{\beta}_{2,2}$. Finally we let $\hat{\sigma}_{2,1-2}^2 = \hat{\sigma}_{2,1}^2 + \hat{\sigma}_{2,2}^2 - 2\hat{\sigma}_{2,12}$ be the estimated variance of $\sqrt{n}(h_{2,1}^T \hat{\beta}_{2,1} - h_{2,2}^T \hat{\beta}_{2,2})$.

We fail to reject the null of $A_2 = 0$ being the best if $-\sqrt{n} h_{2,1}^T \hat{\beta}_{2,1} \geq -d_{2,0} \hat{\sigma}_{2,1}$ and $-\sqrt{n} h_{2,1}^T \hat{\beta}_{2,2} \geq -d_{2,0} \hat{\sigma}_{2,2}$ where $d_{2,0} = d_2(\hat{\sigma}_{2,12}/\hat{\sigma}_{2,1}\hat{\sigma}_{2,2}, \alpha)$ is a data dependent critical value.

To explain d_2 function in detail, $d_2(\rho, \alpha)$ is a deterministic function that takes a scalar ρ satisfying $|\rho| \leq 1$ and another scalar α with $0 \leq \alpha \leq 1$ as inputs and outputs a scalar d. This d satisfies

$$P(\max\{X_1, X_2\} \leq d) = \alpha \tag{3.24}$$

where X_1 and X_2 both follow a $N(0,1)$ distribution and their correlation coefficient

is $\rho.\hat{d}_{2,0}$ is equal to the d satisfying the above relationship when $\rho = \hat{\sigma}_{2,12}/\hat{\sigma}_{2,1}\hat{\sigma}_{2,2}$.

We also provide the procedure of testing whether $A_2 = 1$ and $A_2 = 2$ are the best, respectively. Next we consider treatment $A_2 = 1$. We let 1 into $\hat{S}_2$ if and only if $\sqrt{n}h_{2,1}^T\hat{\beta}_{2,1} \geq -\hat{d}_{2,1}\hat{\sigma}_{2,1}$ and $\sqrt{n}h_{2,1}^T\hat{\beta}_{2,1} - \sqrt{n}h_{2,2}^T\hat{\beta}_{2,2} \geq -\hat{d}_{2,1}\hat{\sigma}_{2,1-2}$ where $\hat{d}_{2,1} = d_2((\hat{\sigma}_{2,1}^2 - \hat{\sigma}_{2,12})/\hat{\sigma}_{2,1}\hat{\sigma}_{2,1-2}, \alpha)$ is a data dependent critical value. Finally we consider treatment $A_2 = 2$. We let 2 into $\hat{S}_2$ if and only if $\sqrt{n}h_{2,2}^T\hat{\beta}_2 \geq -\hat{d}_{2,2}\hat{\sigma}_{2,2}$ and $\sqrt{n}h_{2,2}^T\hat{\beta}_{2,2} - \sqrt{n}h_{2,1}^T\hat{\beta}_{1,1} \geq -\hat{d}_{2,2}\hat{\sigma}_{2,1-2}$, where $\hat{d}_{2,2} = d_2((\hat{\sigma}_{2,2}^2 - \hat{\sigma}_{12})/\hat{\sigma}_{2,2}\hat{\sigma}_{2,1-2}, \alpha)$ is a data dependent critical value.

Remark:

The idea behind this is MCB. Intuitively a treatment cannot be rejected to be the best, if the normalized difference between the estimated effect of this treatment and any other treatment is not too small (here a number is too small means it has large absolute value and a negative sign, not having small absolute value with a positive sign).

Note that $\hat{\sigma}_{2,1}^2(h_2), \hat{\sigma}_{2,2}^2(h_2)$ and $\hat{\sigma}_{2,12}(h_2)$ are all functions of $h_{2,1}, h_{2,2}$ and $\hat{\Sigma}_{2,A}$, where $\hat{\Sigma}_{2,A}$ is the estimated variance matrix of $\sqrt{n}(\hat{\beta}_{2,1}^T, \hat{\beta}_{2,2}^T)^T$ and independent of h_2. As $\hat{d}_{2,0}, \hat{d}_{2,1}, \hat{d}_{2,2}$ are function of those $\hat{\sigma}$s they are all functions of $h_{2,1}, h_{2,2}$ and $\hat{\Sigma}_{2,A}$. So for $H_2 = h_2$, we would write $\hat{S}_2(h_2)$ as

$$\hat{S}_2(h_2) = s_2(h_{2,1}, h_{2,2}, \sqrt{n}\hat{\beta}_{2,1}, \sqrt{n}\hat{\beta}_{2,2}, \hat{\Sigma}_{2,A}, d_2) \tag{3.25}$$

for any $H_2 = h_2$. Here $s_2(\cdot)$ is a deterministic function with range $Power(A_2)$, which is the collection of all subsets of the domain of A_2 except the null set. In our setting for $A_2 \in \{0, 1, 2\}$, $Power(A_2) = \{\{0\}, \{1\}, \{2\}, \{0, 1\}, \{1, 2\}, \{0, 2\}, \{0, 1, 2\}\}$. This notation will be used later.

One note is that $\hat{S}_2$ depends on n only through $\sqrt{n}\hat{\beta}_{2,1}$,$\sqrt{n}\hat{\beta}_{2,2}$ and $\hat{\Sigma}_{2,A}$. Further

more, $\hat{S}_2$ is random only due to the dependence on $\hat{\beta}_2$ and $\hat{\Sigma}_{2,A}$.

As for each second stage treatment, we have probability of including it (in other words, fail to reject the null of it's the best stage-two treatment) if the null is true, no matter which treatment is the true best one, the probability of including it in $\hat{S}_2$ is always no smaller than $1-\alpha$. To be specific, for each $H_2 = h_2$, let

$$[2] = [2](h_2) = \arg\max\{0, h_{2,1}^T \beta_{2,1}^*, h_{2,2}^T \beta_{2,2}^*\}$$

denote the index of the best second stage treatment for patients with $H_2 = h_2$. Let $\hat{S}_2$ be the second stage recommended set constructed following the instruction above, then we have

$$P([2] \in \hat{S}_2) \geq 1 - \alpha$$

for any true β_2^*. This is a immediate result from theorem 1 of *Edwards and Hsu* (1983).

Next we consider the first stage. As has been discussed in the previous sections, as we are providing a set of treatments at the second stage, it is unrealistic to assume a patient will take the true best corresponding second-stage treatment, which is the assumption of the definition using dynamic programming idea. We will allow the patient take the true worst treatment among treatments in his/her recommended set.

To describe this in detail, first we need to define the Q-function for stage one. For any deterministic function s_2 that that maps the domain of H_2 to the power set of $\{0, 1, 2\}$, we define

$$Q_1(h_1, a_1; s_2) := E\big[\min_{a_2 \in s_2(H_2)} Q_2(H_2, a_2) | H_1 = h_1, A_1 = a_1\big] \tag{3.26}$$

This expectation is taken over H_2. Recall that s_2 is a subset of $\{0, 1, 2\}$.

One note is that here H_2 is the only input of s_2 while the s_2 we defined earlier has

many inputs. Later we will show that the s_2 we will use for Q_1 mentioned above is $s_2(h_2) = s_2(h_{2,1}, h_{2,2}, \sqrt{n}\hat{\beta}_{2,1}, \sqrt{n}\hat{\beta}_{2,2}, \hat{\Sigma}_{2,A}, d_2)$. i.e., other inputs are considered fixed. The s_2 mentioned below all mean the s_2 wth H_2 as the only input.

For a given s_2, the parametric form of Q_1 is defined as

$$Q_1(h_1, a_1; \beta_1(s_2), s_2) := h_{1,0}^T \beta_{1,0}(s_2) + h_{1,1}^T \beta_{1,1}(s_2) \mathbb{1}(a_1 = 1) + h_{1,2}^T \beta_{1,2}(s_2) \mathbb{1}(a_1 = 2) \tag{3.27}$$

where $\beta_1(s_2) := (\beta_{1,0}^T(s_2), \beta_{1,1}^T(s_2), \beta_{1,2}^T(s_2))^T$.

$$y(s_2, h_2) := \min_{a_2 \in s_2(h_2)} Q_2(h_2, a_2; \beta_2^*) \tag{3.28}$$

for a patient with h_2. And we denote $\hat{Y}^*(s_2) := y(s_2, H_2)$. Finally, for any second stage treatment regime $\hat{S}_2(h_2) = s_2(h_{2,1}, h_{2,2}, \sqrt{n}\hat{\beta}_{2,1}, \sqrt{n}\hat{\beta}_{2,2}, \hat{\Sigma}_{2,A}, d_2)$, we define our estimand as

$$\hat{\beta}_1^*(\hat{S}_2) = \beta_1(\hat{S}_2) := \left\{ \arg\min_{\beta_1} P(\hat{Y}^*(s_2) - Q_1(H_1, A_1; \beta_1, s_2))^2 \right\} \Big|_{s_2 = \hat{S}_2} \tag{3.29}$$

Remark III.9. $y(s_2, h_2)$ can be interpreted as the expected final outcome for a patient with $H_2 = h_2$ if he or she had taken the worst treatment in $s_2(h_2)$ at the second stage.

Remark III.10. We can see in (3.29) that the estimand, $\hat{\beta}_1^*(\hat{S}_2)$ depends on $\hat{S}_2$, which is a random quantity. This follows the idea of *Robins et al.* (2014). We need to consider the case where decision makers might choose the true worse treatment in the recommended set (if the set contains more than one treatment); we also need to consider the randomness of the recommended set. For notational convenience, we will denote $\hat{\beta}_1^*(\hat{S}_2)$ by $\hat{\beta}_1^*$.

Remark III.11. We can interpret $\hat{\beta}_1^*$ as the true first-stage regression coefficient if the second-stage recommended set is $\hat{S}_2$. At stage one for a new patient with $H_1 = h_1$, under the assumption that the second-stage recommended set is $\hat{S}_2$, the patient's

expected final outcome is $h_{1,0}^T\hat{\beta}_{1,0}^* + h_{1,1}^T\hat{\beta}_{1,1}^*\mathbb{1}(A_1 = 1) + h_{1,2}^T\hat{\beta}_{1,2}^*\mathbb{1}(A_1 = 2)$. Thus we are interested in the vector $(h_{1,1}^T\hat{\beta}_{1,1}^*, h_{1,2}^T\hat{\beta}_{1,2}^*)$.

Remark III.12. To help understand our procedure, we are using one data set to mimic the following procedure based on data sets from TWO clinical trials: from the first trial $\hat{S}_2$ is constructed. Next consider the second trial, in which we use the mapping $\hat{S}_2$ to offer treatment at each patient's second stage. Then for a patient with $H_1 = h_1$, this patient's expected final outcome is $h_{1,0}^T\hat{\beta}_{1,0}^*+h_{1,1}^T\hat{\beta}_{1,1}^*\mathbb{1}(A_1 = 1)+h_{1,2}^T\hat{\beta}_{1,2}^*\mathbb{1}(A_1 = 2)$. Note that this expected final outcome is conditional on the $\hat{S}_2$ from the first data set. But of course we only have one data set, so our procedure mimics the above but with only one data set.

Remark III.13. In stage one, we construct the first stage recommended set $\hat{S}_1$ similar to how we do in stage two. First, for every patient, construct their hypothetical final outcome

$$\tilde{Y}(\hat{S}_2; H_2) := \min_{a_2 \in \hat{S}_2} (H_{2,0}^T\hat{\beta}_{2,0} + H_{2,1}^T\hat{\beta}_{2,1}\mathbb{1}(a_2 = 1) + H_{2,2}^T\hat{\beta}_{2,2}\mathbb{1}(a_2 = 2)). \tag{3.30}$$

as an estimator of $y(s_2, h_2)$ in (3.28) with $s_2 = \hat{S}_2$. Next we get the estimator of $\hat{\beta}_1^*$ in (3.29) by

$$\hat{\beta}_1 := \arg\min_{\beta_1} P_n(\tilde{Y} - (H_{1,0}^T\beta_{1,0} + H_{1,1}^T\beta_{1,1}\mathbb{1}(A_1 = 1) + H_{1,2}^T\beta_{1,2}\mathbb{1}(A_1 = 2)))^2 \tag{3.31}$$

Before we move on to discuss in detail how to construct $\hat{S}_1$ we introduce some notation. For $s \subset \{0, 1, 2\}$, we denote

$$F(s, a, b, c) := \min_{A \in s}(a + b\mathbb{1}(A = 1) + c\mathbb{1}(A = 2)) \tag{3.32}$$

Thus (3.30) can be expressed as $\tilde{Y} = H_{2,0}^T\hat{\beta}_{2,0} + F(\hat{S}_2, 0, H_{2,1}^T\hat{\beta}_{2,1}, H_{2,2}^T\hat{\beta}_{2,2})$ and (3.28) can be expressed as $\hat{Y}^* = H_{2,0}^T\beta_{2,0}^* + F(\hat{S}_2, 0, H_{2,1}^T\beta_{2,1}^*, H_{2,2}^T\beta_{2,2}^*)$.

For stage one, we consider the three first stage treatments one by one. For $A_1 = 0$, we are interested in a two dimensional vector, each component of which is the difference between the effect of $A_1 = 0$ and $A_1 = 1, 2$. To be specific we are interested in $(-H_{1,1}^T\hat{\beta}_{1,1}^*, -H_{1,2}^T\hat{\beta}_{1,2}^*)^T$. Similarly for $A_1 = 1$ and $A_1 = 2$ we are interested in $(H_{1,1}^T\hat{\beta}_{1,1}^*, H_{1,1}^T\hat{\beta}_{1,1}^* - H_{1,2}^T\hat{\beta}_{1,2}^*)^T$ and $(H_{1,2}^T\hat{\beta}_{1,2}^*, H_{1,2}^T\hat{\beta}_{1,2}^* - H_{1,1}^T\hat{\beta}_{1,1}^*)^T$. We include an index in our first stage recommended set $\hat{S}_1$ if and only if we fail to reject the null that both components of the vector of interested are non-negative. Thus, for $A_1 = 0$, we would like to construct a simultaneous one sided 1-α level confidence intervals for $(-H_{1,1}^T\hat{\beta}_{1,1}^*, -H_{1,2}^T\hat{\beta}_{1,2}^*)^T$, say $((-\infty, u_1), (-\infty, u_2))^T$, (meaning $P(-H_{1,1}^T\hat{\beta}_{1,1}^* \leq u_1, -H_{1,2}^T\hat{\beta}_{1,2}^* \leq u_2) \geq 1 - \alpha$), then if both u_1 and u_2 are non-negative we fail to reject the null at $1 - \alpha$ level and we would let 0 into $\hat{S}_1$. The test statistic here is naturally $(-H_{1,1}^T\hat{\beta}_{1,1}, -H_{1,2}^T\hat{\beta}_{1,2})$. The procedure of whether letting $A_1 = 1, 2$ into $\hat{S}_1$ are similar.

To make the problem more general, we are interested in constructing one sided simultaneous confidence intervals for $(c_1\hat{\beta}_1^*, c_2\hat{\beta}_1^*)^T$ where c_1 and c_2 are two row vectors with length equal to the length of β_1. We are interested in the vector $(\sqrt{n}c_1(\hat{\beta}_1 - \hat{\beta}_1^*), \sqrt{n}c_2(\hat{\beta}_1 - \hat{\beta}_1^*))^T$. Let B_1 be the first stage covariate $B_1 = (H_{1,0}^T, H_{1,1}^T\mathbb{1}(A_1 = 1), H_{1,2}^T\mathbb{1}(A_1 = 2))^T$ and $\hat{\Sigma}_1 := \mathbb{P}_n B_1 B_1^T$ then we have $\hat{\beta}_1 = \hat{\Sigma}_1^{-1}\mathbb{P}_n B_1\tilde{Y}$, similarly define $\hat{\beta}_1^* = \Sigma_{1,\infty}^{-1} P B_1 \hat{Y}^*$ where $\Sigma_{1,\infty} := P B_1 B_1^T$. Thus, we can show that $\sqrt{n}c_1(\hat{\beta}_1 - \hat{\beta}_1^*) = c_1(\mathbb{S}_n + \mathbb{U}_n)$ where

$$\begin{aligned}
\mathbb{S}_n &= \sqrt{n}\hat{\Sigma}_1^{-1}\mathbb{P}_n B_1\Big[H_{2,0}^T\beta_{2,0}^* + F(\hat{S}_2, 0, H_{2,1}^T\beta_{2,1}^*, H_{2,2}^T\beta_{2,2}^*) - B_1\hat{\beta}_1^* + H_{2,0}^T(\hat{\beta}_{2,0} - \beta_{2,0}^*)\Big] \\
\mathbb{U}_n &= \sqrt{n}\hat{\Sigma}_1^{-1}\mathbb{P}_n B_1\Big[F(\hat{S}_2, 0, H_{2,1}^T\hat{\beta}_{2,1}, H_{2,2}^T\hat{\beta}_{2,2}) - F(\hat{S}_2, 0, H_{2,1}^T\beta_{2,1}^*, H_{2,2}^T\beta_{2,2}^*)\Big]
\end{aligned}$$

Below are some remarks:

1. Term $\mathbb{S}_n$ is always asymptotically normal. (proof will be included in the

appendix)

2. Term $\mathbb{U}_n$, on the other hand, is not asymptotically normal when there are some equally best second stage treatments for a certain kind of patients. For example for some $h_{2,1}$ and $h_{2,2}$ that satisfy both $P(H_{2,1} = h_{2,1}, H_{2,2} = h_{2,2}) > 0$ and $h_{2,1}^T \beta_{2,1}^* = 0 > h_{2,2}^T \beta_{2,2}^*$ (this is the case that at the second stage for some patients treatment 0 and 1 have the same effect while treatment 2 is the worst, other scenarios are similar), when n is large with probability converging to one we will not include $A_2 = 2$ into $\hat{S}_2$. Thus we have

$$\begin{aligned}
& \sqrt{n}(F(\hat{S}_2, 0, H_{2,1}^T \hat{\beta}_{2,1}, H_{2,2}^T \hat{\beta}_{2,2}) - F(\hat{S}_2, 0, H_{2,1}^T \beta_{2,1}^*, H_{2,2}^T \beta_{2,2}^*)) \\
= & \sqrt{n} F(\hat{S}_2, 0, H_{2,1}^T \hat{\beta}_{2,1}, H_{2,2}^T \hat{\beta}_{2,2}) \\
= & \sqrt{n}(\min(0, H_{2,1}^T \hat{\beta}_{2,1}) \mathbb{1}(\hat{S}_2 = \{0,1\}) + H_{2,1}^T \hat{\beta}_{2,1} \mathbb{1}(\hat{S}_2 = \{1\})) \\
= & \sqrt{n}(-[H_{2,1}^T \hat{\beta}_{2,1}]_- \mathbb{1}(\hat{S}_2 = \{0,1\}) + H_{2,1}^T \hat{\beta}_{2,1} \mathbb{1}(\hat{S}_2 = \{1\}))
\end{aligned}$$

where $[x]_-$ equals to 0 if $x \geq 0$ and $-x$ if $x < 0$. We can see that this term is not asymptotically normal. Thus the distribution of $\mathbb{U}_n$ and its bootstrap analog $\mathbb{U}_n^{(b)}$ will have different limiting distribution. Thus the confidence interval constructed from bootstrap method will have poor coverage rate.

Now we discuss the lower bound of $(c_1\sqrt{n}(\hat{\beta}_1 - \hat{\beta}_1^*), c_2\sqrt{n}(\hat{\beta}_1 - \hat{\beta}_1^*))^T$ (Note that then we can use this lower bound to form the upper bound of $(c_1\hat{\beta}_1^*, c_2\hat{\beta}_1^*)$) . From the above discussion, we already have that

$$\begin{pmatrix} c_1\sqrt{n}(\hat{\beta}_1 - \hat{\beta}_1^*) \\ c_2\sqrt{n}(\hat{\beta}_1 - \hat{\beta}_1^*) \end{pmatrix} = \begin{pmatrix} c_1(\mathbb{S}_n + \mathbb{U}_n) \\ c_2(\mathbb{S}_n + \mathbb{U}_n) \end{pmatrix} \tag{3.33}$$

What we want is an lower bound for the two dimensional vector (3.33), which means we would like a data dependent value l such that the probability of both

components of (3.33) above l is no smaller than $1-\alpha$, i.e.,

$$P(c_1\sqrt{n}(\hat{\beta}_1-\hat{\beta}_1^*)\geq l, c_2\sqrt{n}(\hat{\beta}_1-\hat{\beta}_1^*)\geq l)\geq 1-\alpha \tag{3.34}$$

then after some algebra we have $P(c_1\hat{\beta}_1^*\leq c_1\hat{\beta}_1-l/\sqrt{n}, c_2\hat{\beta}_1^*\leq c_2\hat{\beta}_1-l/\sqrt{n})\geq 1-\alpha$.

The idea of ACI (Adaptive Confidence Interval) is that we split our data into two parts, those we determine the unique best second stage treatment and those we cannot. For the former part, we just use normal bootstrap method since we are quite sure that there is no "non-regularity" (details of how the bootstrap method is applied is introduced in Tianshuang's former paper); while for the latter part, we use γ_1 and γ_2 to replace $\sqrt{n}\beta_{2,1}^*$ and $\sqrt{n}\beta_{2,2}^*$ and let them vary in their domain. Before we introduce the explicit form of the lower bound we denote $\mathbb{V}_n=\sqrt{n}((\hat{\beta}_{2,1}^T,\hat{\beta}_{2,2}^T)^T-(\beta_{2,1}^{*T},\beta_{2,2}^{*T})^T)$, $\mathbb{V}_{n,1}=\sqrt{n}(\hat{\beta}_{2,1}-\beta_{2,1}^*)$ and $\mathbb{V}_{n,2}=\sqrt{n}(\hat{\beta}_{2,2}-\beta_{2,2}^*)$. We also define $\tilde{S}_2(\lambda_n)=\hat{S}_2(H_{2,1},H_{2,2},\sqrt{n}\hat{\beta}_{2,1},\sqrt{n}\hat{\beta}_{2,2},\lambda_n)$. Note that λ_n is not a function like d_2 but a constant (we can also treat λ_n as a constant function). As we will see we will let λ_n go to infinity with rate slower than $\sqrt{n}$ to ensure that the indices included in $\tilde{S}_2$ are eventually all the indices of the true best second stage treatments. Finally, note that we can write the left hand side of (3.34) in one dimension form as $P(\min\big(c_1\sqrt{n}(\hat{\beta}_1-\hat{\beta}_1^*), c_2\sqrt{n}(\hat{\beta}_1-\hat{\beta}_1^*)\big)\geq l)$. A lower bound of $\min\big(c_1\sqrt{n}(\hat{\beta}_1-\hat{\beta}_1^*), c_2\sqrt{n}(\hat{\beta}_1-\hat{\beta}_1^*)\big)$ is given by

$$\mathcal{L}(c_1,c_2)=\inf_{\gamma_1,\gamma_2}\big(\min\{\mathcal{F}(c_1,\gamma_1,\gamma_2),\mathcal{F}(c_2,\gamma_1,\gamma_2)\}\big) \tag{3.35}$$

where

$$
\begin{aligned}
& \mathcal{F}(c,\gamma_1,\gamma_2) \\
= \;& c^T\mathbb{S}_n + c^T\mathbb{U}_n\mathbb{1}(|\tilde{S}_2(\lambda_n)|=1) \\
& +\Bigg[c^T\hat{\Sigma}_1^{-1}\mathbb{P}_n B_1\Big[F\big(\hat{S}_2(H_{2,1},H_{2,2},\gamma_1+\mathbb{V}_{n,1},\gamma_2+\mathbb{V}_{n,2},d_2),0,H_{2,1}^T(\gamma_1+\mathbb{V}_{n,1}), \\
& H_{2,2}^T(\gamma_2+\mathbb{V}_{n,2})\big) \\
& -F\big(\hat{S}_2(H_{2,1},H_{2,2},\gamma_1+\mathbb{V}_{n,1},\gamma_2+\mathbb{V}_{n,2},d_2),0,H_{2,1}^T\gamma_1,H_{2,2}^T\gamma_2\big)\Big]\Bigg]\mathbb{1}(|\tilde{S}_2(\lambda_n)|>1)
\end{aligned}
$$

Here $|S|$ means the number of components of the set S.

Similarly, although we will not use it in this paper, an upper bound of $\max\big(c_1\sqrt{n}(\hat{\beta}_1-\hat{\beta}_1^*),c_2\sqrt{n}(\hat{\beta}_1-\hat{\beta}_1^*)\big)$ is given by

$$
\mathcal{U}(c_1,c_2)=\sup_{\gamma_1,\gamma_2}\big(\max\{\mathcal{F}(c_1,\gamma_1,\gamma_2),\mathcal{F}(c_2,\gamma_1,\gamma_2)\}\big) \tag{3.36}
$$

As we have $c_i\sqrt{n}(\hat{\beta}_1-\hat{\beta}_1^*)=\mathcal{F}(c_i,\sqrt{n}\beta_{2,1}^*,\sqrt{n}\beta_{2,2}^*)$ for $i=1,2$, so we have $\mathcal{L}(c_1,c_2)\le\min\big(c_1\sqrt{n}(\hat{\beta}_1-\hat{\beta}_1^*),c_2\sqrt{n}(\hat{\beta}_1-\hat{\beta}_1^*)\big)$ and $\max\big(c_1\sqrt{n}(\hat{\beta}_1-\hat{\beta}_1^*),c_2\sqrt{n}(\hat{\beta}_1-\hat{\beta}_1^*)\big)\le\mathcal{U}(c_1,c_2)$ from the following lemma

Lemma III.14. *Let γ be a vector and F_1,F_2 two functions of γ, we have*

$$
\min\{\inf_\gamma F_1(\gamma),\inf_\gamma F_2(\gamma)\}=\inf_\gamma\min\{F_1(\gamma),F_2(\gamma)\}
$$

Proof. $\forall\gamma$, $\min\{F_1(\gamma),F_2(\gamma)\}\ge LHS$. Take infimum of γ we have $RHS\ge LHS$.

$\inf_\gamma F_1(\gamma)\ge RHS$ and $\inf_\gamma F_2(\gamma)\ge RHS$ together imply $LHS\ge RHS$. So we finish our proof. □

Remark III.15. In standard MCB procedure, instead of finding lower bound of $\min\big(c_1\sqrt{n}(\hat{\beta}_1-\hat{\beta}_1^*),c_2\sqrt{n}(\hat{\beta}_1-\hat{\beta}_1^*)\big)$, we would like to find a lower bound of $\min\big(c_1\sqrt{n}(\hat{\beta}_1-\hat{\beta}_1^*)/std(c_1\sqrt{n}(\hat{\beta}_1-\hat{\beta}_1^*)),c_2\sqrt{n}(\hat{\beta}_1-\hat{\beta}_1^*)/std(c_2\sqrt{n}(\hat{\beta}_1-\hat{\beta}_1^*))\big)$ where

std mean the standard deviation. We don't do it here because it is very hard to estimate $std(c\sqrt{n}(\hat{\beta}_1 - \hat{\beta}_1^*))$. Our procedure is still valid, in the sense that we would still include the true best first stage treatment in the recommended set if we use this form. Proof can be found in the later section. And if we assume that the variance of each treatment effect estimator are close then our procedure is the same as standard MCB.

Define

$$
\begin{aligned}
&\mathbb{W}_\infty = \\
&\Sigma_{1,\infty}^{-1}\mathbb{G}B_1\Big[H_{2,0}^T\beta_{2,0}^* + \\
&F(\hat{S}_2(H_{2,1}, H_{2,2}, \sqrt{n}\beta_{2,1}^* + \mathbb{V}_1, \sqrt{n}\beta_{2,2}^* + \mathbb{V}_2, \Sigma_{2,A,\infty}, d_2), 0, H_{2,1}^T\beta_{2,1}^*, H_{2,2}^T\beta_{2,2}^*) \\
&-B_1^T\Sigma_{1,\infty}^{-1}PB_1\big[H_{2,0}^T\beta_{2,0}^* + \\
&F(\hat{S}_2(H_{2,1}, H_{2,2}, \sqrt{n}\beta_{2,1}^* + \mathbb{V}_1, \sqrt{n}\beta_{2,2}^* + \mathbb{V}_2, \Sigma_{2,A,\infty}, d_2), 0, H_{2,1}^T\beta_{2,1}^*, H_{2,2}^T\beta_{2,2}^*)\big] \\
&+H_{2,0}^T\mathbb{V}_0\Big]
\end{aligned}
\tag{3.37}
$$

We denote $\mathbb{V}_{n,i} = \sqrt{n}(\hat{\beta}_{2,i} - \beta_{2,i}^*)$ for $i = 0.1, 2$, and $\mathbb{V}_i$ the limiting distribution correspondingly. We have

$$
\begin{aligned}
&\sqrt{n}c^T(\hat{\beta}_1 - \hat{\beta}_1^*) \\
\to\ & c^T\mathbb{W}_n + c^T\Sigma_{1,\infty}^{-1}PB_1\Big[\\
&F(\hat{S}_2(H_{2,1}, H_{2,2}, \sqrt{n}\beta_{2,1}^* + \mathbb{V}_1, \sqrt{n}\beta_{2,2}^* + \mathbb{V}_2, \Sigma_{2,A,\infty}, d_2), 0, H_{2,1}^T(\sqrt{n}\beta_{2,1}^* + \mathbb{V}_1), \\
&H_{2,2}^T(\sqrt{n}\beta_{2,2}^* + \mathbb{V}_2)) \\
&-F(\hat{S}_2(H_{2,1}, H_{2,2}, \sqrt{n}\beta_{2,1}^* + \mathbb{V}_1, \sqrt{n}\beta_{2,2}^* + \mathbb{V}_2, \Sigma_{2,A,\infty}, d_2), 0, \sqrt{n}H_{2,1}^T\beta_{2,1}^*, \\
&\sqrt{n}H_{2,2}^T\beta_{2,2}^*)\Big]
\end{aligned}
$$

Denote $S^* = S^*(H_2) = \{i \in \{0,1,2\} | i = \arg\max_i H_{2,1}^T\beta_{2,1}^*\mathbb{1}(i=1) + H_{2,2}^T\beta_{2,2}^*\mathbb{1}(i =$

2)} be the set of indices of the best second stage treatment for patient with H_2.

The limiting distribution of $\inf_{\gamma_1,\gamma_2} \mathcal{F}(c,\gamma_1,\gamma_2)$ is given by

$$
\begin{aligned}
&c^T\mathbb{W}_\infty + c^T\Sigma_{1,\infty}PB_1\Big[H_{2,0}^T\mathbb{V}_0 + H_{2,1}^T\mathbb{V}_1\mathbb{1}(H_{2,1}^*\beta_{2,1}^* > \max\{0, H_{2,2}^T\beta_{2,2}^*\}) \\
&+H_{2,2}^T\mathbb{V}_2\mathbb{1}(H_{2,2}^*\beta_{2,2}^* > \max\{0, H_{2,1}^T\beta_{2,1}^*\})\Big]\mathbb{1}(|S^*|=1) \\
&+\inf_{\gamma_1,\gamma_2}\Bigg\{c^T\Sigma_{1,\infty}^{-1}PB_1\Big[\\
&F(\hat{S}_2(H_{2,1},H_{2,2},\gamma_1+\mathbb{V}_1,\gamma_2+\mathbb{V}_2,d_2),0,H_{2,1}^T(\gamma_1+\mathbb{V}_1),H_{2,2}^T(\gamma_2+\mathbb{V}_2)) \\
&-F(\hat{S}_2(H_{2,1},H_{2,2},\gamma_1+\mathbb{V}_1,\gamma_2+\mathbb{V}_2,d_2),0,H_{2,1}^T\gamma_1,H_{2,2}^T\gamma_2)\Big]\mathbb{1}(|S^*|>1)\Bigg\}
\end{aligned}
$$

Thus we can guarantee the probability of including $\min\big(c_1\sqrt{n}(\hat{\beta}_1-\hat{\beta}_1^*), c_2\sqrt{n}(\hat{\beta}_1-\hat{\beta}_1^*)\big)$ in the one-sided confidence interval.

3.7.3 Simulation Study

Similar to section 3.4, we would like to examine the performance of our methods. Again we propose some generative models with different "degrees of non-regularity". We use the following generative model

$$
\begin{aligned}
Y \;=\;& \gamma_1+\gamma_2X_1+\gamma_3\mathbb{1}(A_1=1)+\gamma_4\mathbb{1}(A_1=2)+\gamma_5X_1\mathbb{1}(A_1=1)+\gamma_6X_1\mathbb{1}(A_1=2) \\
&+\big(\gamma_7+\gamma_8X_1+\gamma_9\mathbb{1}(A_1=1)+\gamma_{10}\mathbb{1}(A_1=2)\big)\mathbb{1}(A_2=1) \\
&+\big(\gamma_{11}+\gamma_{12}X_1+\gamma_{13}\mathbb{1}(A_1=1)+\gamma_{14}\mathbb{1}(A_1=2)\big)\mathbb{1}(A_2=2)+\epsilon
\end{aligned}
$$

Below are some information about the simulation.

- Each trajectory has the form (X_1, A_1, X_2, A_2, Y).
- For each sample, its A_1 and A_2 are assigned with probability 1/3 to be one of 0,1,2.
- X_1 is the initial condition, with 1/2 probability of being -1 and 1.

- X_2 is the second stage covariate, as we will not use it for analyzing we will assign the value of it with the same distribution as X_1. It is here only to make the trajectory similar to the two-treatments=per-stage case.

- ϵ follows iid $N(0,1)$ distribution.

- $H_{2,0} = (1, X_1, \mathbb{1}(A_1 = 1), \mathbb{1}(A_1 = 2), X_1\mathbb{1}(A_1 = 1), X_1\mathbb{1}(A_1 = 2))$

- $H_{2,1} = H_{2,2} = (1, X_1, \mathbb{1}(A_1 = 1), \mathbb{1}(A_1 = 2))$

- $H_{1,0} = H_{1,1} = H_{1,2} = (1, X_1)$.

- $\alpha = 0.05$. i.e., we want to include the best first stage treatment with probability no smaller than 0.95.

We will focus on the case where $\gamma_9 = \gamma_{10} = \gamma_{13} = \gamma_{14}$, in this case if patients starts with $A_1 = 0$, all the three second stage treatments will be equally the best for them. On the other hand, if they start with $A_1 = 1$ or $A_1 = 2$, only one of the three second stage treatment is the best, and the effect is equal to the effects starting with $A_1 = 0$. In this case the effect of $A_1 = 0$ will be highly underestimated. In this case, the ACI method can cover the best treatment $A_1 = 0$ with probability 0.951 while the naive bootstrap method's probability is only 0.83.

3.7.4 Discussion

We want to further examine if there is any less conservative method, since when all the treatments have equal effect, the ACI method will include the true best first stage treatment with probability very closed to one.

CHAPTER IV

Identifying a set that contains the best DTR

The work presented in this chapter (except that in section 4.8) is joint work with Ashkan Ertefaie who was a postdoc in our group when working on this, Kevin Lynch who is a researcher in Pennsylvania, and Inbal Nahum-Shani who is an assistant professor in the Institute for Social Research of University of Michigan.

The work in section 4.8 is a joint with Rong Zhou who is an undergrad student in the department of Statistics.

4.1 Introduction

A dynamic treatment regime (DTR) is a treatment design that seeks to accommodate patient heterogeneity in response to treatments (*Murphy et al.*, 2001; *Murphy*, 2003; *Robins*, 2004). In DTRs the type and/or dose of the treatment is adapted over time according to the patient's characteristics and progress in treatment. At each decision point (i.e., specific point in time in which a treatment is to be considered or altered), decision rules are used to map individual characteristics to a specific type of treatment or dosage. Recently, there has been an increased interest in sequential, multiple assignment, randomized trials (SMARTs), which were developed specifically to provide empirical evidence that informs the construction of optimal DTRs (*Lavori and Dawson*, 2000; *Laber et al.*, 2014b; *Nahum-Shani et al.*, 2012a; *Chakraborty and*

Moodie, 2013; *Chakraborty and Murphy*, 2014).

One scientific question motivating a SMART concerns the comparison of DTRs that are embedded in the design. It aims to identify the best DTR or the set that contains the best DTRs among those that are embedded in the design. In other words, the goal is to *screen out* ineffective DTRs. This question can be framed as a special case of the general multiple comparison problem.

Methods for multiple comparisons can be used to group sample means, such that within each group, population means are not significantly different (*Tukey*, 1953; *Scheffe*, 1953). Current approaches for identifying the set of best DTRs perform all possible comparisons among embedded DTRs. In such a setting, standard multiple comparison approaches used to control for Type I error result in a loss of statistical power (*Hsu et al.*, 1984; *Hsu*, 1996). Consequently, important differences between DTRs might go undetected (*Saville*, 1990; *Keselman et al.*, 1999). Here, we propose a more efficient approach for identifying the set of best DTRs. This approach builds on the work of *Hsu et al.* (1981), which identifies the best set of means by conducting multiple comparisons with the best (MCB), namely by comparing the best mean with others. Applying this approach will result in fewer comparisons relative to standard approaches, and hence improved power.

The current manuscript will extend the MCB toolbox for analyzing data from SMART studies. The contribution of this paper is twofold. First, we provide and illustrate, for the first time, a method that can be used to efficiently address an important scientific question that motivates many SMART studies. This question concerns the need to identify the optimal DTR, or several optimal DTRs from a list of DTRs embedded in a SMART study. Enabling researchers to address this scientific question can support clinical decision making, offering clinicians a set of efficacious DTRs to choose from based on other considerations such as cost and patient preferences. The second contribution concerns the correlation structure of the estimators derived from

SMART data. The method proposed by Hsu requires a known correlation structure (up to a constant). In SMART, the correlation structure of estimators are not known a priori. Therefore, generalization of the method os warranted.

We briefly introduce SMART designs and explain the structure of SMART data in Section 4.2. We then present two methods to estimate the mean outcome under each DTR in Section 4.3. The framework of MCB in SMART settings is introduced in Section 4.4. We conduct a simulation study in Section 4.5 to examine the performance of our method. We illustrate the method with analyses of the Extending Treatment Effectiveness of Naltrexone (EXTEND) study in Section 4.6. The last section contains some concluding remarks. Proofs are given in an online supplementary document.

4.2 Preliminaries

4.2.1 Sequential, Multiple Assignment, Randomized Trials

The SMART is a clinical trial design in which each individual proceeds through stages of treatments (*Lavori et al.*, 2000; *Murphy*, 2005a; *Lei et al.*, 2012; *Nahum-Shani et al.*, 2012a). At each treatment stage, individuals are randomized to one of the available treatment options at that stage, where the subsequent treatment options may depend on an embedded tailoring variable observed at current stage. For example, in the EXTEND study, at stage 1, patients were randomized to one of two definitions of non-response while receiving naltrexone (NTX): 1) Stringent criterion– a patient is a non-responder if (s)he has two or more heavy drinking days in the first eight weeks; 2) Lenient criterion– a patient is a non-responder if (s)he has five or more heavy drinking days in the first eight weeks. At stage 2, non-responders were re-randomized to combined behavioral intervention (CBI) +NTX or CBI alone. Individuals who did not meet their non-response criterion were re-randomized to telephone disease management (TDM)+NTX or NTX alone. Thus, in this two-stage

design, the embedded tailoring variable is the response/non-response status to initial NTX.

4.2.2 Data Structure

For simplicity, we focus on SMARTs with two stages. The observed data on each individual are given by a trajectory $(O_1, A_1, O_2, S, A_2, Y)$. O_j, for $j = 1, 2$ is a set of covariates available at the beginning of stage j. A_j denotes the treatment options at the beginning of stage j. S is a binary variable that is coded 1 if an individual has been re-randomized at stage 2, and coded 0 otherwise. Finally, Y is the continuos primary outcome. The treatment and the covariate history through j are denoted by $\bar{A}_j$ and $\bar{O}_j$, respectively. We use lowercase letters to refer to the possible values of the corresponding capital letter random variable.

In SMART settings, the stage-2 treatment options may depend on embedded tailoring variables, which are part or all of the observed history up to and including time 2, and we denote them as V. In the EXTEND study, V is the response(R)/non-response(NR) status to stage-1 treatment (i.e., $V \in \{R, NR\}$). Hence, for each individual, we conceptualize a v-treatment trajectory $\mathcal{T} = (A_1, A_2^R, A_2^{NR})$. For responders and non-responders, we set $A_2^{NR} = 0$ and $A_2^R = 0$, respectively, with probability 1. This basically means that for responders A_2^{NR} does not apply and vice versa. We use the v-treatment trajectory to model the marginal structural model discussed in Section 3. Note, the v-treatment trajectory and treatment history are not necessarily the same. In fact, in this example, the treatment history is two-dimensional, while the v-treatment trajectory is three-dimensional.

4.2.3 Embedded Dynamic Treatment Regimes

An embedded dynamic treatment regime (EDTR) is one DTR that participants can follow as part of the study design. In the EXTEND study, there are 8 EDTRs:

1) Start with lenient definition. If the patient is non-responsive, offer NTX+CBI; if the patient is responsive, offer NTX+TDM. 2) Start with lenient definition. If the patient is non-responsive, offer NTX+CBI; if the patient is responsive, offer NTX. 3) Start with lenient definition. If the patient is non-responsive, offer CBI; if the patient is responsive, offer NTX+TDM. 4) Start with lenient definition. If the patient is non-responsive, offer CBI; if the patient is responsive, offer NTX. The other four EDTRs are similar except that they start with stringent definition. Note that a given v-treatment trajectory $\mathcal{T}$ can be consistent with more than one EDTR. For example, a responder to the lenient definition with $\mathcal{T} = (A_1 = \text{lenient}, A_2^R = NTX + TDM, A_2^{NR} = 0)$ is following both EDTRs (1) and (3).

4.3 Estimation

Let θ_k be the population outcome mean under the kth EDTR for $k = 1, 2, ..., K$ where K is the number of EDTRs in a SMART. Here, we provide two methods that are based on weighted least squares minimizations and used throughout this paper as tools to estimate the mean outcome under each EDTR. The first approach would be to postulate a marginal structural model (MSM) $m(\mathcal{T}; \beta_{p\times 1})$ for the outcome given the observed v-treatment trajectory $\mathcal{T}$ and define θ_k as a known function of β for all k. Let $\mathbb{P}_n$ be the empirical average. The parameters of the MSM can be estimated using the following estimating equation:

$$\mathbb{P}_n \sum_{k=1}^{K} \dot{m}(\mathcal{T}; \beta) w_2(V, \bar{A}_2, k)(Y - m(\mathcal{T}; \beta)) = 0, \tag{4.1}$$

where $\dot{m}(\mathcal{T}; \beta) = \partial m(\mathcal{T}; \beta)/\partial\beta$, and

$$w_2(v, \bar{a}_2, k) = \frac{I_{EDTR_{k,1}}(a_1) I_{EDTR^v_{k,2}}(a_2)}{p(A_1 = a_1) p(A_2 = a_2 | A_1 = a_1, V = v)}, \quad \text{for } V = v \text{ and } \bar{A}_2 = \bar{a}_2,$$

where $EDTR^V_{k,2}$ is the treatment option determined by EDTR$_k$ at stage 2 given V, and $EDTR_{k,1}$ is the treatment option determined by EDTR$_k$ at stage 1. The indicator function selects individuals whose treatment history is consistent with the kth EDTR given V. This method is referred to as inverse probability weighting (IPW) (*Robins*, 1999; *Robins et al.*, 2000; *Hernán et al.*, 2000). The treatment trajectory is used to define the MSM function. For example, in the EXTEND study, the MSM would be $m(\mathcal{T};\beta) = \beta_0 + \beta_1 A_1 + \beta_2 A_2^R + \beta_3 A_2^{NR} + \beta_4 A_1 A_2^R + \beta_5 A_1 A_2^{NR}$. We denote the solutions of this equation as $\hat{\beta}^{IPW}$. Hence, the mean outcome under each EDTR can be estimated as $(\hat{\theta}_1^{IPW}, ..., \hat{\theta}_K^{IPW}) = D\hat{\beta}^{IPW}$, where D is a $K \times p$ matrix. The kth row of D is the contrast corresponding to EDTR$_k$ (see Section 4.5).

The second approach is based on the augmented IPW (AIPW), which is a more efficient version of IPW developed by *Robins et al.* (2008) and *Orellana et al.* (2010). Let $EDTR^V_k = (EDTR_{k,1}, EDTR^V_{k,2})$. The corresponding estimating equation for a two-stage design is given by

$$\begin{aligned}\mathbb{P}_n \sum_{k=1}^{K} \dot{m}(\mathcal{T};\beta)\Big[& w_2(V,\bar{A}_2,k)(Y - m(\mathcal{T};\beta)) \\ & - \big(w_2(V,\bar{A}_2,k) - w_1(A_1,k)\big)\big(\varphi_2^k(\bar{O}_2) - m(\mathcal{T};\beta)\big) \\ & - \big(w_1(A_1,k) - 1\big)\big(\varphi_1^k(O_1) - m(\mathcal{T};\beta)\big)\Big] = 0, \qquad (4.2)\end{aligned}$$

where $\varphi_2^k(\bar{O}_2) = \mathbb{E}[Y|\bar{A}_2 = EDTR^V_k, \bar{O}_2]$, $\varphi_1^k(O_1) = \mathbb{E}[\varphi_2^k(\bar{O}_2)|A_1 = EDTR_{k,1}, O_1]$, and

$$w_1(a_1,k) = \frac{I_{EDTR_{k,1}(a_1)}}{p(A_1 = a_1)}, \text{ for } A_1 = a_1.$$

To obtain estimators of β, we postulate parametric models for the unknown functions $\varphi_1^k(.)$ and $\varphi_2^k(.)$ parametrized by γ and replace them with their estimated values $\varphi_1^k(.,\hat{\gamma})$ and $\varphi_2^k(.,\hat{\gamma})$. The estimates may be obtained by fitting two least squares models. We denote the solutions of (4.2) as $\hat{\beta}^{AIPW}$ and, similar to the first approach,

we define $(\hat{\theta}_1^{AIPW}, ..., \hat{\theta}_K^{AIPW}) = D\hat{\beta}^{AIPW}$.

Estimator (4.2) is double robust in the sense that it results in an unbiased estimate of β if either $\varphi_k(., \gamma)$ or the treatment assignment probabilities are correctly specified (*van der Laan and Robins*, 2003; *Davidian et al.*, 2005; *Bang and Robins*, 2005; *Orellana et al.*, 2010). Although we are focusing on randomized trials and treatment assignment probabilities are known by design, for efficiency we estimate these probabilities nonparametrically using the available data (*Robins et al.*, 1995; *Hirano et al.*, 2003). One may also postulate a parametric model to estimate these probabilities given the observed covariate/treatment history.

The following proposition provides the asymptotic behaviours of estimators $\hat{\theta}^{IPW}$ and $\hat{\theta}^{AIPW}$ obtained by (4.1) and (4.2), respectively, which is an immediate consequence of Lemma 3 in *Orellana et al.* (2010). In the proposition, the superscript $\diamond$ denotes IPW or AIPW.

Proposition IV.1. Let $\hat{\theta}^{\diamond} = D\hat{\beta}^{\diamond}$, where D is a $K \times p$ matrix with the kth row of D being the contrast corresponding to the kth EDTR. Then under the standard regularity assumptions, $\sqrt{n}(\hat{\theta}^{\diamond} - \theta) \to N\left(0, \Sigma^{\diamond} = D'[\Gamma'^{-1}\Lambda^{\diamond}\Gamma^{-1}]D\right)$, where $\Gamma = -\mathbb{E}\left[\sum_{i=1}^{K} \dot{m}'(\mathcal{T}; \beta)\dot{m}(\mathcal{T}; \beta)\right]$, and $\Lambda^{\diamond} = \mathbb{E}[U'^{\diamond}U^{\diamond}]$ with

$$
\begin{aligned}
U^{AIPW} = &\sum_{k=1}^{K} \dot{m}(\mathcal{T}; \beta)\Big[w_2(V, \bar{A}_2, k)(y - m(\mathcal{T}; \beta)) \\
&- \big(w_2(V, \bar{A}_2, k) - w_1(A_1, k)\big)\big(\varphi_2^k(\bar{O}_2) - m(\mathcal{T}; \beta)\big) \\
&- \big(w_1(A_1, k) - 1\big)\big(\varphi_1^k(O_1) - m(\mathcal{T}; \beta)\big)\Big], \\
U^{IPW} = &\sum_{k=1}^{K} \dot{m}(\mathcal{T}; \beta)\Big[w_2(V, \bar{A}_2, k)(y - m(\mathcal{T}; \beta))\Big].
\end{aligned}
$$

The asymptotic variance $\Sigma^{\diamond}$ may be estimated consistently by replacing the expectations with expectations with respect to the empirical measure and (β, γ) with its estimate $(\hat{\beta}^{\diamond}, \hat{\gamma})$ and denoted as $\hat{\Sigma}^{\diamond} = D[\hat{\Gamma}^{-1}\hat{\Lambda}^{\diamond}\hat{\Gamma}^{-1}]D$.

4.4 Multiple Comparison with the Best

Let $\mathcal{B}$ be the true set of best EDTRs and $\hat{\mathcal{B}}$ be a set of EDTRs that cannot be differentiated from the best EDTR using the available data. In the previous section, we discussed our procedures to estimate the mean outcome under each EDTR, $\hat{\theta}_k^{IPW}$ and $\hat{\theta}_k^{AIPW}$ for $k = 1, 2, ..., K$. Since our methodology holds for both IPW and AIPW approaches to estimation, for simplicity of notation, we drop the superscripts IPW and AIPW and refer to the estimator of θ_k as $\hat{\theta}_k$. In this section, we generalize the MCB method introduced by *Hsu et al.* (1981) to SMART settings. The goal is to find EDTRs that are *not* significantly different from the EDTR with the maximum outcome, say $\theta_{[K]} = max_{1 \leq k \leq K} \theta_k$. Hence, a natural criterion would be to include index i in the set $\hat{\mathcal{B}}$ if the standardized difference $(\hat{\theta}_i - \hat{\theta}_j)/\sigma_{ij}$ is greater than a constant for all $j \neq i$. This can be written as

$$\hat{\theta}_i \geq max_{j \neq i}[\hat{\theta}_j - c_i \sigma_{ij}], \tag{4.3}$$

where c_i is a constant and $\sigma_{ij} = \sqrt{var(\hat{\theta}_i - \hat{\theta}_j)}$, which can be estimated using the variance formula in Proposition IV.1. The challenge is to find c_i such that $\hat{\mathcal{B}}$ includes the true best EDTR with probability at least $(1 - \alpha)$; that is, $P_\theta(\arg\max_i \theta_i \in \hat{\mathcal{B}}) \geq 1 - \alpha$ for any θ. In cases where there are more than one best EDTR, $\hat{\mathcal{B}}$ includes each index $k \in \mathcal{B}$ with at least $(1 - \alpha)$ probability. This condition will be satisfied if we find c_i such that under the null hypothesis (i.e., all EDTRs are equally good), the set $\hat{\mathcal{B}}$ includes each index k with probability $(1 - \alpha)$. In other words, when σ_{ij} is known, c_i must satisfy

$$p(Z_i \geq Z_j - c_i \sigma_{ij} \text{ for } j = 1, 2, i-1, i+1, ..., K) = 1 - \alpha, \tag{4.4}$$

where $Z_1, ..., Z_K$ are multivariate normal random variables with mean 0 and covariance matrix Σ such that $\sqrt{var(Z_i - Z_j)} = \sigma_{ij}$. The above equality can be written as

$$\int p(Z_1 \leq z+c_i\sigma_{ij}, ..., Z_{i-1} \leq z+c_i\sigma_{ij}, Z_{i+1} \leq z+c_i\sigma_{ij}, ..., Z_K \leq z+c_i\sigma_{ij})d\phi(z) = 1-\alpha,$$

where $\phi(z)$ is the marginal cdf of Z_i. Note that for $\alpha \leq 0.5$, the constant $c_i > 0$ for $i = 1, ...K$. Hence, in our setting where α represents the Type I error rate, we can assume that c_i is a positive constant.

Hsu et al. (1981) present an equation that can be used to find the constant c_i when the structure of the covariance matrix Σ is known up to a constant. This is the case in a standard regression where $\Sigma = \sigma(X'X)^{-1}$. Note that in this case, given the design matrix, Σ is known up to a constant σ. In Hsu's setting, the constant c_i is a function of the correlation matrix and thus it is not a function of σ. In the marginal structural model, however, the structure of the design matrix is random because it depends on intermediate outcomes (i.e., variables observed before stage 2 and after stage 1 treatment assignment) that are not included in the design matrix, such as response or non-response status (i.e., embedded tailoring variables). In such setting, the constant c_i will be a function of an unknown Σ which is estimated by $\hat{\Sigma}$ using the observed data. Theorem IV.2 generalizes the idea in Hsu to cases where the structure of the design matrix is unknown. We use the notation $\hat{c}_i$ to reflect the dependence of c_i to $\hat{\Sigma}$.

Theorem IV.2. *Define the estimated set of best EDTRs as* $\hat{\mathcal{B}} = \{i|\hat{\theta}_i \geq max_{j\neq i}[\hat{\theta}_j - \hat{c}_i\hat{\sigma}_{ij}]\}$, *where* $\hat{\sigma}_{ij} = \sqrt{\hat{v}ar(\hat{\theta}_i - \hat{\theta}_j)}$, *and* $\hat{c}_i$ *satisfies*

$$\int p(Z_1 \leq z + \hat{c}_i\hat{\sigma}_{i1}, ..., Z_{i-1} \leq z + \hat{c}_i\hat{\sigma}_{i(i-1)},$$
$$Z_{i+1} \leq z + \hat{c}_i\hat{\sigma}_{i(i+1)}, ..., Z_K \leq z + \hat{c}_i\hat{\sigma}_{iK})d\phi(z) = 1 - \alpha,$$

with $Z_1, ..., Z_K$ *being multivariate normal random variables with mean 0 and unknown*

covariance matrix Σ*, which is estimated by* $\hat{\Sigma}$*. Then, asymptotically,* $\hat{\mathcal{B}}$ *contains the true best EDTR with probability at least* $(1-\alpha)$*.*

Let $\delta_{ij} = \theta_i - \theta_j$ be the difference between the ith and the jth EDTR. The probability of including an EDTR$_i$ in the estimated set of best EDTRs for any given c_i is

$$
\begin{aligned}
&p(EDTR_i \in \hat{\mathcal{B}}) \\
&= p\left(\hat{\theta}_i \geq \hat{\theta}_1 - c_i\hat{\sigma}_{i1}, ..., \hat{\theta}_i \geq \hat{\theta}_{i-1} - c_i\hat{\sigma}_{i(i-1)}, \hat{\theta}_i \geq \hat{\theta}_{i+1} - c_i\hat{\sigma}_{i(i+1)}, ..., \hat{\theta}_i \geq \hat{\theta}_K - c_i\hat{\sigma}_{i(K)}\right) \\
&= p\Big(W_{i1} \leq c_i\frac{\hat{\sigma}_{i1}}{\sigma_{i1}} + \frac{\delta_{i1}}{\sigma_{i1}}, ..., W_{i(i-1)} \leq c_i\frac{\hat{\sigma}_{i(i-1)}}{\sigma_{i(i-1)}} + \frac{\delta_{i(i-1)}}{\sigma_{i(i-1)}}, W_{i(i+1)} \leq c_i\frac{\hat{\sigma}_{i(i+1)}}{\sigma_{i(i+1)}} + \frac{\delta_{i(i+1)}}{\sigma_{i(i+1)}} \\
&\qquad\qquad , ..., W_{iK} \leq c_i\frac{\hat{\sigma}_{iK}}{\sigma_{iK}} + \frac{\delta_{iK}}{\sigma_{iK}}\Big), \qquad (4.5)
\end{aligned}
$$

where $W_{ij} = -(\hat{\theta}_i - \hat{\theta}_j - \delta_{ij})/\sigma_{ij}$ and is distributed as a standard normal random variable. Accordingly, the estimated set size (ESS) of $\hat{\mathcal{B}}$ is defined as $\sum_{i=1}^{K} p(EDTR_i \in \hat{\mathcal{B}})$. Note, under the null hypothesis, where all EDTRs are equally good, $ESS = K(1-\alpha)$. The following theorem shows that the probability of including an inferior $EDTR_i$ in the estimated set $\hat{\mathcal{B}}$ decays to zero exponentially for $i = 1, ..., K$ as the difference between the best and the ith EDTR increases.

Theorem IV.3. *For any fixed index* i*, let* $W_i = (W_{i1}, \cdots, W_{iK})$ *follow a multivariate normal distribution with mean zero and unknown variance matrix. Define* $Y_j = c_i\frac{\hat{\sigma}_{ij}}{\sigma_{ij}}$ *for* $j = 1, ..., i-1, i+1, ..., K$ *as non-negative random variables. (Note that* Y_j *also depends on* i *but as we are fixing* i*, for notation convenience we omit* i *here). Let* $K = argmax_{1\leq i\leq K}\theta_i$*,* $\forall\delta_{Ki} \geq 0$*; we have*

$$P\left(W_{i1} \leq Y_1, W_{i2} \leq Y_2, \cdots, W_{iK} \leq Y_K - \frac{\delta_{Ki}}{\sigma_{iK}}\right)$$
$$\leq P\left(W_{i1} \leq Y_1, W_{i2} \leq Y_2, \cdots, W_{iK} \leq Y_K\right) \exp\left(-\zeta \delta_{Ki}/\sigma_{iK}\right), \tag{4.6}$$

where $\zeta > 0$ is a constant that depends on $cov(W_i)$ and $Y_1, \cdots, Y_K$ but is independent of δ_{Ki}.

Note that $\sigma_{iK} = \sqrt{var(\hat{\theta}_i - \hat{\theta}_K)}$, which decays to zero with rate $1/\sqrt{n}$. This implies that for a fixed δ_{Ki}, as n increases the probability of including an inferior $EDTR_i$ to $\hat{\mathcal{B}}$ decreases with rate $\exp(-\sqrt{n})$. Also, in the statement of Theorem IV.3, replacing Y_i with $Y_i + \frac{\delta_{i1}}{\sigma_{i1}}$, for $i = 1, ..., K-1$, shows the exponential decay rate in (4.5).

Remark. Let Σ^{AIPW} and Σ^{IPW} be the covariance matrix of $cov(\hat{\theta}^{AIPW})$ and $cov(\hat{\theta}^{IPW})$, respectively. Since $\Sigma^{AIPW} \preceq \Sigma^{IPW}$, for any fixed sample size and a set of δ_{ij}s, the efficient estimator AIPW results in an ESS which is less than or equal to the one obtained by the inefficient estimator IPW (see Figures A.1 and A.2).

4.5 Simulation Study

This section provides empirical evidence for the theoretical results presented in the manuscript. We compare the estimated sets of best obtained by the IPW and AIPW methods and show that the latter method screens out the ineffective EDTRs more efficiently. We examine the performance of the proposed method using two different types of SMART designs. We describe that the form of the marginal structural model $m(., \beta)$ may vary based on the design structure. We also discuss the effect that misspecifying the function $\varphi^k_{\cdot}(.)$ has on estimating the parameters of the marginal structural model and the mean outcome under each EDTR.

In all simulation scenarios, baseline variables O_{11} and O_{12} are generated from standard normal, and $A_1 \in \{-1, +1\}$ is based on a Bernoulli distribution with probability 0.5. The intermediate outcomes are $O_{21} \sim N(0.5O_{11}, 1)$ and $O_{22} \sim N(0.5O_{12}, 1)$. The estimator IPW and AIPW refer to (4.1) and (4.2), respectively, while AIPW$_m$ refers to an AIPW estimator where $\varphi^k_{\cdot}(.)$ functions are misspecified. All the *Web* tables and figures corresponding to this section are presented in Appendix A.4 of the supplementary material available at *Biometrics* online.

4.5.1 SMART Design: Example 1

This is a type of SMART design in which just a subset of individuals are rerandomized at stage 2. In our simulation, this subset is non-responders to stage-1 treatment (see Web Figure 1). Thus, the embedded tailoring variable $V \in \{R, NR\}$ is the indicator of responder or non-responder status, respectively. Four DTRs are embedded in this design depending on v-treatment trajectory $\mathcal{T} = (A_1, A_2^{NR})$; these are listed in Web Table 1. Note that because there is only one treatment option for responders, the v-treatment trajectory does not include A_2^R. We generate these SMART data with sample sizes 100, 200, 300 and 400 from the following generative model. The stage-2 treatment option $A_2^{NR} \in \{-1, +1\}$ is generated from a Bernoulli distribution with probability 0.5. The outcome is generated from normal distribution with mean $1 + O_{11} - O_{12} + O_{21} + O_{22} + A_1(\delta + O_{11}) + S\delta/2A_2^{NR}$, with variance $\sigma^2 = 1$, where $S = I(O_{21} > 0)$. The main effect of treatment options are parametrized with δ. The true $\varphi^k_{\cdot}$s are given by

$$\varphi_2^k(\bar{o}_2, \bar{a}_2 \in EDTR_k^V, \gamma) = \gamma_0 + \gamma_1 o_{11} + \gamma_2 o_{12} + \gamma_3 o_{21} + \gamma_4 o_{22} + a_1(\gamma_5 + \gamma_6 o_{11}) + \gamma_7 s a_2,$$

$$\varphi_1^k(o_1, a_1 \in EDTR_k^V, \gamma) = \gamma_8 + \gamma_9 o_{11} + \gamma_{10} o_{12} + \gamma_{11} a_1 + \gamma_{12} a_1 o_{11}.$$

We also consider a misspecified scenario where $\varphi_2^k(\bar{o}_2, \bar{a}_2 = EDTR_k^V, \gamma^\dagger) = \gamma_0^\dagger + \gamma_1^\dagger o_{11} + \gamma_2^\dagger o_{21} + \gamma_3^\dagger o_{22}$ and $\varphi_1^k(o_1, a_1 = EDTR_{k,1}, \gamma^\dagger) = \gamma_4^\dagger + \gamma_5^\dagger o_{11}$ are assumed to be working models. Moreover, the marginal structural model is

$$m(\mathcal{T}; \beta) = \beta_0 + \beta_1 A_1 + \beta_2 A_2^{NR}.$$

Hence, the true parameter value $\beta^* = (1.00, \delta, \delta/4)$, which means, for $\delta > 0$, $(A_1 = +1, A_2^{NR} = +1)$ is the true best EDTR and, for $\delta < 0$, $(A_1 = -1, A_2^{NR} = -1)$ is the true best EDTR. Table A.1 presents the point estimate and standard errors of the parameters $\beta_{3\times1}$ and $\theta_{4\times1} = D\beta$ estimated using IPW, AIPW and AIPW$_m$, where

$$D = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \end{pmatrix}.$$

The rows of this matrix represent $EDTR_1, ..., EDTR_4$ listed in Web Table 1. In Table A.1, we set $\delta = 0.1$ and generated 1000 datasets of sizes 100 and 400. Our results show that AIPW reduces the standard error by up to 60% compared to IPW, and even when $\varphi_\cdot^k(.)$ functions are misspecified AIPW$_m$ maintains unbiasedness, but some of the standard errors increase. In fact, under our misspecification scenario AIPW still has better performance than IPW. We see a similar pattern in estimation of the mean outcome under different EDTRs.

Figure A.1 shows how fast the size of the set of best $\hat{\mathcal{B}}$ converges to 1 as δ increases when the parameters of each EDTR is estimated using IPW and AIPW. Note that for $\delta \neq 0$, the true set size $\mathcal{B}$ is 1. For each δ, we generated 500 data sets and defined the ESS as the empirical average of the set sizes for each data set. This figure shows that when the parameters β of the marginal structural model are estimated using AIPW, the ESS decreases to 1 faster than when using IPW. This is due to the more

efficient estimation of βs.

4.5.2 SMART Design: Example 2

In some SMART designs, stage-2 randomization depends on prior treatment *and* an intermediate outcome such as response indicator (see Web Figure 2). We generate datasets of sizes 100, 200, 300 and 400 from the following generative model. The stage-2 treatment options are generated from a multinomial distribution with probability 0.25 coded as 1,2,3, and 4. Let $R \in \{0, 1\}$ be the non-response and response indicator, respectively. Then, if $V = I(A_1 = +1) + I(A_1 = -1)I(R = 1) = 1$ (i.e., V satisfies condition A), there is no randomization; while individuals with $V = 0$ (i.e., V satisfies condition B) will be randomized to one of the four stage-2 treatment options. Hence the v-treatment trajectory in this example is $\mathcal{T} = (A_1, A_2^B)$. Five DTRs are embedded in this design depending on the treatment trajectory (A_1, A_2^B); these are listed in Web Table 2.

The outcome is generated from a normal distribution with mean $1 + O_{11} - O_{12} + O_{21} + O_{22} + I(A_1 = -1)(\delta + O_{11}) + SI(A_1 = -1)[-\delta/4I(A_2 = 1) + \delta/2I(A_2 = 2) + 0I(A_2 = 3) + \delta/2O_{21}I(A_2 = 2)]$ with variance $\sigma^2 = 1$, where $S = I(O_{21} > 0)$. Thus the true $\varphi_\cdot^k(.)$s are

$$
\begin{aligned}
&\varphi_2^k(\bar{o}_2, \bar{a}_2 = EDTR_k^V, \gamma) \\
&= \gamma_0 + \gamma_1 o_{11} + \gamma_2 o_{12} + \gamma_3 o_{21} + \gamma_4 o_{22} + I(a_1 = -1)(\gamma_5 + \gamma_6 o_{11}) \\
&+ sI(a_1 = -1)[\gamma_7 I(a_2 = 1) + \gamma_8 I(a_2 = 2) + \gamma_9 I(a_2 = 3) + \gamma_{10} o_{21} I(a_2 = 2)], \\
&\varphi_1^k(o_1, a_1 \in EDTR_{k,1}, \gamma) = \gamma_{11} + \gamma_{12} o_{11} + \gamma_{13} o_{12} + \gamma_{14} I(a_1 = -1) + \gamma_{15} I(a_1 = -1) o_{11}.
\end{aligned}
$$

We also consider a misspecified scenario where $\varphi_2^k(\bar{o}_2, \bar{a}_2 = EDTR_k^V, \gamma^\dagger) = \gamma_0^\dagger + \gamma_1^\dagger o_{11} + \gamma_2^\dagger o_{21} + \gamma_3^\dagger o_{22} + \gamma_4^\dagger a_1 + sI(a_1 = -1)[\gamma_5^\dagger I(a_2 = 1) + \gamma_6^\dagger I(a_2 = 2) + \gamma_7^\dagger I(a_2 = 3)]$ and $\varphi_1^k(o_1, a_1 = EDTR_{k,1}, \gamma^\dagger) = \gamma_8^\dagger + \gamma_9^\dagger o_{11} + \gamma_{10}^\dagger a_1$ are assumed to be working models.

The MSM is

$$m(\mathcal{T};\beta) = \beta_0 + \beta_1 I(A_1 = -1) + I(A_1 = -1)[\beta_2 I(A_2^B = 1) + \beta_3 I(A_2^B = 2) + \beta_4 I(A_2^B = 3)].$$

Hence, the true parameter value $\beta^* = (1.00, \delta, -\delta/8, \delta/4, 0)$, which means that for positive and negative δs, $(A_1 = -1, A_2^B = 2)$ and $(A_1 = -1, A_2^B = 1)$ are the best EDTRs, respectively. Table A.2 presents the bias and standard errors of the parameters $\beta_{5\times1}$ and $\theta_{5\times1} = D\beta$ estimated using IPW, AIPW and AIPW$_m$, where

$$D = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

The rows of this matrix represent $EDTR_1, ..., EDTR_5$ listed in Web Table 2. In Table A.2, we set $\delta = 0.4$ and generated 1000 datasets of sizes 100 and 400. Our results show that AIPW reduces the standard error of θs and βs by up to 55% compared to IPW. The misspecified scenario, where the interaction terms in both $\varphi_\cdot^k(.)$ functions are ignored, results in estimators with slightly larger standard errors compared to $AIPW$.

Figure A.2 shows how fast the size of the set of best converges to 1 as δ grows when the parameters of each EDTR are estimated using IPW and AIPW. Note that for $\delta \neq 0$ the true set size $\mathcal{B}$ is 1. For each δ, we generated 500 datasets and defined the ESS as the empirical average of the set sizes for each data set. This figure shows that when the parameters β of the marginal model are estimated using AIPW, the estimated set size decreases to 1 faster than when using IPW. This is due to more efficient estimation of βs. The plot of ESS when estimated using AIPW$_m$ is omitted

since it is similar to IPW in this simulation.

4.6 Illustrative data analysis

The EXTEND study was a 24-week, multistage clinical trial that enrolled alcohol dependent patients (*Lei et al.*, 2012). At stage 1, patients are randomized with probability 0.5 to either the stringent or lenient definitions of non-response while receiving naltrexone (NTX). Participants were assessed weekly for drinking behavior, and starting at week 3, as soon as the participant met his/her assigned criterion for early non-response, he/she was immediately re-randomized to one of two rescue tactics: (1) offering CBI in addition to NTX (i.e., NTX+CBI); or (2) offering CBI alone (i.e., CBI). Participants who did not meet their assigned criterion for early non-response by the end of week 8 (i.e., responders to NTX) were re-randomized at that point (i.e., end of week 8) to one of two maintenance tactics: either (1) adding TDM to NTX (i.e., NTX+TDM) or offering NTX alone (NTX). Web Figure 3 (Appendix A.4 of the supplementary material) depicts this two-stage SMART design.

For illustration we focus on a simplified version of this trial. Let the primary outcome Y denote the Penn Alcohol Craving Scale (PACS) score over 24 weeks. Lower PACSs are preferable. Let A_1 denote the non-response criterion coded as -1 for stringent and +1 for lenient. The embedded tailoring variable V in this design is the response/non-response status. The stage-2 treatment options for responders are NTX ($A_2^R = -1$) and NTX+TDM ($A_2^R = +1$) and for non-responders the rescue treatment options are CBI ($A_2^{NR} = -1$) and NTX+CBI ($A_2^{NR} = +1$). Additionally, let R denote the indicator for whether (R=1) or not (R=0) the patient was a responder to the initial NTX treatment. Web Figure 3 in Appendix A.4 of the supplementary material available at *Biometrics* online shows the number of patients assigned to each treatment option. By design, there are 2^3 EDTRs in this SMART based on different combinations of (A_1, A_2^R, A_2^{NR}), which are listed in Web Table 3.

Baseline variables include PACS before stage 1 (O_{11}) and gender (O_{12}). The intermediate outcomes are the average PACS during stage 1 (O_{21}) and the standard error of the measured PACS during stage 1 (O_{22}). We consider the following marginal structural model: $m(\mathcal{T}, \beta) = \beta_0 + \beta_1 A_1 + \beta_2 A_2^R + \beta_3 A_2^{NR}$. One may add the interaction terms $A_1 A_2^R$ and $A_1 A_2^{NR}$ to this model. Also, we consider $\varphi_2^k(\bar{o}_2, \bar{a}_2 = EDTR_k^V, \gamma) = \gamma_0 + \gamma_1 o_{11} + \gamma_2 o_{12} + \gamma_3 o_{21} + \gamma_4 o_{22} + a_1(\gamma_5 + \gamma_6 o_{11}) + r a_2(\gamma_7 + \gamma_8 o_{21}) + (1-r) a_2(\gamma_9 + \gamma_{10} o_{21})$ and $\varphi_1^k(o_1, a_1 = EDTR_{k,1}, \gamma) = \gamma_{11} + \gamma_{12} o_{11} + \gamma_{13} o_{12} + \gamma_{14} a_1 + \gamma_{15} a_1 o_{11} + \gamma_{16} a_1 o_{12}$.

We estimated the parameter vector $\beta_{4\times 1}$ and $\theta_{8\times 1} = D\beta$ using both the IPW (4.1) and AIPW (4.2) estimators and the results are presented in Table A.3, where

$$D = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & -1 \end{pmatrix}.$$

The rows of this matrix represent $EDTR_1, ..., EDTR_8$ listed in Web Table 3. The point estimate and standard errors for β_0 and β_1 are very close using both estimators. However, the parameters corresponding to A_2^R and A_2^{NR} have smaller standard errors when estimated using AIPW. Moreover, our procedure screens out EDTR_6 and EDTR_8 when the parameter vector β is estimated using AIPW, but using IPW results in keeping all eight EDTRs in the set of best. In other words, when using MCB with the AIPW approach to estimate the mean outcome under each EDTR, results indicated that DTRs that begin with NTX, classifies patients as non-responders by using a stringent criterion, and offers CBI alone to non-responders and NTX or NTX+TDM to responders, do not belong to the set of best EDTRs.

4.7 Discussion

An important research question motivating many SMART studies concern the selection of the best (i.e., most efficacious) DTR among a set of DTRs that are embedded in a SMART. However, this is often not possible due to small sample size. In this manuscript, we propose a method that can be used to identify the *set* that contains the best DTR. We frame the problem as a special case of multiple comparison and show that the constructed set of best contains the true best DTR with at least a given probability. We use the AIPW estimator to estimate the mean under each DTR, and our simulation results show that for any given sample size the cardinality of the constructed set of best is less than the cardinality obtained by IPW estimators, while maintaining the Type I error rate. Moreover, we prove that the probability of inclusion of an inferior DTR in the constructed set of best decays exponentially as the difference between the best and the inferior DTR grows.

Currently most SMART designs are sized such that an investigator can detect either a given stage-1 or stage-2 treatment effect or a given difference between two DTRs with a given probability. One important extension of this work would be to devise a method that can be used to plan SMART sample sizes such that the constructed set of best includes at most m DTRs, for a fixed difference between the best and the worst DTRs, with a given probability. This will be more consistent with the goal of SMART designs in many applications.

4.8 Comparison with the modified version of ACI method

The work presented in this section is a joint work with undergrad student Rong Zhou.

Another way of constructing a set that contains the best DTR is the modified version of the ACI method discussed in chapter III. But there are two aspects that

are different from that method. Below we will introduce this modified method in detail.

First, we will adapt the dynamic programming idea discussed in *Laber et al.* (2010), which is different from our idea in chapter III. Recall that we argue that we assume that patients will choose the true worst treatment in the future while *Laber et al.* (2010) assumes that patients will choose the true best treatment in the future.

The second difference is that instead of treating each treatment as a unit, we will treat a DTR as a unit. Thus the we will try to evaluate the "effect" of a DTR. Below is the detailed procedure of the modified ACI method. One remark is that this modified ACI method can only be applied to embedded DTRs.

1. For the second stage treatments, if we have more than one options, form the recommended set $\hat{S}_2(A_1, R)$ where A_1 is the first stage treatment and R is the response indicator. The procedure of forming the set is the same as that in chapter III

2. Define the hypothetical outcome $\tilde{Y}$ and form the confidence interval for the "effect" of a first stage treatment as in *Laber et al.* (2010).

3. Use the adaptive confidence interval to construct the first stage recommended set $\hat{S}_1$.

4. A DTR is included in the recommended set of DTR if and only if each of its compartment is in the corresponding recommended set of treatments.

As an example of the last step, consider a set of eight DTRs where at the first stage, at the second stage for responders and at the second stage for non-responders, there are two available treatments. We use a triplet (A_1, A_{2R}, A_{2NR}) to denote a DTR where each of the three can take values 1 and -1. For a particular DTR (a_1, a_{2R}, a_{2NR}) to be included in the recommended set of DTRs, we need all $a_1 \in \hat{S}_1$, $a_{2R} \in \hat{S}_2(A_1 = a_1, R = 1)$ and $a_{2NR} \in \hat{S}_2(A_1 = a_1, R = 0)$ to be true.

Another remark is that since we are doing three hypothesis testing for one DTR, for now we use Bonferroni correction to control the overall error rate. This means if

we want to include the true best DTR with probability $1-\alpha$, then for each test we will use $1-\alpha/3$ as the significance level.

We would like to check the performance of these two methods: the one introduced in this chapter and denoted as MCB method, and the modified ACI method which is denoted as ACI method for short. We will use the same generative model as in chapter III as follows:

$$Y = \gamma_1 + \gamma_2 X_1 + \gamma_3 A_1 + \gamma_4 X_1 A_1 + \gamma_5 A_2 + \gamma_6 X_2 A_2 + \gamma_7 A_1 A_2 + \epsilon, \epsilon \sim N(0,1)$$

For details of how A_1, R and A_2 are generated please refer to the simulation section in chapter III.

We will compare the two methods in four scenarios. Each scenario represents a case where a certain "degree of non-regularity" is obtained. We will compare the probability of containing the true best DTR as well as the average set size. The latter is a measure of how conservative a method is. On average, subjective to containing the true best DTR with at least the given probability, a more conservative method will end up in a set with larger average set size. Naturally the size of a set is the number of DTRs it contains.

Intuitively, MCB method will perform better when the degree of non-regularity if high, as in this case all DTRs are close and this is the least conservative scenario for MCB, while in this case ACI will be the most conservative because the second stage recommended set of treatments will contain both treatments with high probability, ending up with a conservative confidence interval for the difference of effects of the two first stage treatments. On the other hand, when the model is "regular", i.e., there is a significant difference between each pair of second stage treatments, ACI method will be the least conservative while MCB method will be conservative since the number of DTRs to be compared is unnecessarily large, resulting in an unnecessarily large

recommended set of DTRs.

The detailed description and discussion are included in Appendix A.5

CHAPTER V

Discussion and future work

5.1 Discussion

In this thesis we discussed several techniques and application of DTRs. The main focus is to construct a recommended set of DTRs or treatments at a certain stage. This technique allows clinicians and patients to have more choices according to their own preferences. The treatments or DTRs contained in the recommended set are those we cannot differentiate from the best while those we excluded are what we believe to be inferior. The goal of constructing the set is not to pin down the sole best treatment or DTR, but to exclude inferior options.

In order to construct the recommended set, we need to compare treatments or DTRs, thus we review and develop many comparisons methods in this thesis. Each of the methods discussed has its own advantages and disadvantages, and researchers will choose among them according to their preference and the scientific questions they are trying to answer.

The main part of this thesis is chapter III, where we introduce a new definition of the effect of a non-final stage treatment. Following this definition, we construct the recommended set of non-final stage treatments using ACI technique to overcome the non-regularity problems. A consistency theorem is proved as well. Simulation studies shows both the poor behavior of the naive bootstrap method under some certain

scenario and the power of our new method. This method is generalized to the case where we have more than two treatments per stage.

5.2 Future work

There is a long way along the set valued policy setting. First, we have not proved the consistency result for the case where there are more than two treatments per stage. This will be a main focus after my graduation. Also, we are using the "un-standardized" vector in the hypothesis test. This will be fine if the true variances of the estimator of the effects of the treatments are close, which should be the common case in reality. However we need to examine the behavior of our method under the scenario that the variances differ by a considerably large amount.

Another direction, as has been discussed at the end of chapter III, is that we will try to extend our procedure to the case where there are more than three stages. *Laber et al.* (2010) discussed this situation but they assumed that the covariance of the estimated difference between a certain treatment and the true best treatment, is the covariance of the estimated difference between this treatment and the treatment with the larges point estimator. Mathematically they assumed $\widehat{var}(\hat{\theta}_i - \hat{\theta}_{[N]}) = \widehat{var}(\hat{\theta}_i - \hat{\theta}_{(N)})$ where $[N] = \arg\max_i \theta_i$ and $(N) = \arg\max_i \hat{\theta}_i$. In our future work we will try to examine how bad it can be if this is violated.

APPENDIX

APPENDIX A

Related Proof and other supplements

A.1 The proof of theorem III.5

This section provides the proof of all the theorems in section 3. The proof of these theorems requests several lemmas. First, we state some results without proving. They are results for the second stage estimation and the proofs are exactly the same as those in *Laber et al.* (2010).

Theorem A.1. *Assume (A1) and (A2) and fix* $a \in \mathbb{R}^{p_2}$*, then*

1. $a^T\sqrt{n}(\hat{\theta}_2 - \theta_2^*) \xrightarrow{d}_P a\mathbb{Z}_\infty$*,*

2. $a^T\sqrt{n}(\hat{\theta}_2^{(b)} - \hat{\theta}_2) \xrightarrow{d}_{P_M} a\mathbb{Z}_\infty$ *in P-probability; and*

3. if in addition (A3) holds, $a^T\sqrt{n}(\hat{\theta}_2 - \theta_{2,n}^*) \xrightarrow{d}_{P_M} a\mathbb{Z}_\infty$

where $\mathbb{Z}_\infty$ *is a mean zero normal random vector with covariance matrix* $\Sigma_{2,\infty}^{-1}P[B_2B_2^T(Y_2 - B_2^T\theta_2^*)^2]\Sigma_{2,\infty}^{-1}$.

Theorem A.2. *Assume (A1),(A2), then* $\hat{\Sigma}_2 \to_P \Sigma_{2,\infty}$*, and* $\hat{\Sigma}_2^{(b)} \to_{P_M} \Sigma_{2,\infty}$ *in P-probability as* $n \to \infty$*. Furthermore if (A3) holds then* $\hat{\Sigma}_2 \to_{P_n} \Sigma_{2,\infty}$ *as* $n \to \infty$*.*

Now we prove the main theorem. First we define a sequence of functions to have a

decomposition of $\mathcal{U}(c)$.

1. $\omega_{11} : D_{p_1} \times D_{p_1 \times p_{20}} \times D_{p_{21}} \times l^\infty(\mathcal{F}_{11}) \times l^\infty(\mathcal{F}_{11}) \times \mathbb{R}^{p_2} \times \mathbb{R}^{p_{21}} \times \mathbb{R}^{p_1+p_2} \to \mathbb{R}$ is defined as

$$
\begin{aligned}
& \omega_{11}(\Sigma_1, \Sigma_{12}, \Sigma_{21,21}\mu, \omega, \nu, \eta, \theta) \\
:= \;& \mu(c^T\Sigma_1^{-1}B_1(Y_1 + H_{2,1}^T\theta_{2,0} - |H_{2,1}^T\theta_{2,1}|\mathbb{1}(T(H_{2,1}, \eta+\nu_1, \Sigma_{21,21}) \le \chi) \\
& + H_{2,1}^T\theta_{2,1} sign(H_{2,1}^T(\eta+\nu_1))\mathbb{1}(T(H_{2,1}, \eta+\nu_1, \Sigma_{21,21}) > \chi) - B_1\theta_1) \\
& + c^T\Sigma_1^{-1}\Sigma_{12}\nu_0 + \omega(c^T\Sigma_1^{-1}B_1H_{2,1}^T\nu_1\mathbb{1}_{H_{2,1}^T\theta_{2,1}^*>0}) - \omega(c^T\Sigma_1^{-1}B_1H_{2,1}^T\nu_1\mathbb{1}_{H_{2,1}^T\theta_{2,1}^*<0})
\end{aligned}
$$

where $\mathcal{F}_{11} = \{f(b_1, y_1, h_{2,0}, h_{2,1}) = a_1^Tb_1(y_1 + h_{2,0}^T\theta_{2,0} - |h_{2,1}^T\theta_{2,1}|\mathbb{1}(T(h_{2,1}, \eta+\nu_1, \Sigma_{21,21}) \le \chi) + h_{2,1}^T\theta_{2,1}sign(h_{2,1}^T(\eta+\nu_1))\mathbb{1}(T(h_{2,1}, \eta+\nu_1, \Sigma_{21,21}) > \chi) - b_1^T\theta_1) + a_2^Tb_1(h_{2,1}^T\nu_1)\mathbb{1}_{h_{2,1}^T\theta_{2,1}^*>0} - a_2^Tb_1(h_{2,1}^T\nu_1)\mathbb{1}_{h_{2,1}^T\theta_{2,1}^*<0} : \theta = (\theta_1^T, \theta_{2,0}^T, \theta_{2,1}^T)^T \in \mathbb{R}^{p_1+p_2}, \nu = (\nu_0^T, \nu_1^T)^T \in \mathbb{R}^{p_2}, \eta \in \mathbb{R}^{p_{21}}.a_1.a_2 \in \mathbb{R}^{p_1}, \max\{||a_1||, ||a_2||, ||\theta||, ||\nu||, ||\eta||\} \le K\}$.

2. $\omega_{12} : D_{p_1} \times D_{p_{21}} \times l^\infty(\mathcal{F}_{12}) \times \mathbb{R}^{p_{21}} \times \mathbb{R}^{p_{21}} \to \mathbb{R}$ is defined as

$$
\begin{aligned}
\omega_{12}(\Sigma_1, \Sigma_{21,21}, \mu, \nu, \gamma) \;:=\; & \mu\Big[c^T\Sigma_1^{-1}B_1(|H_{2,1}^T\nu + H_{2,1}^T\gamma| - |H_{2,1}^T\gamma|) \\
& \times(2\mathbb{1}(T(H_{2,1}, \nu+\gamma, \Sigma_{21,21}) > \chi) - 1))\mathbb{1}_{H_{2,1}^T\theta_{2,1}^*=0}\Big]
\end{aligned}
$$

where $\mathcal{F}_{12} = \{f(b_1, h_{2,1}) = a^Tb_1(|h_{2,1}^T(\nu+\gamma)| - |h_{2,1}^T\gamma|)(2\mathbb{1}(T(h_{2,1}, \nu+\gamma, \Sigma_{21,21}) > \chi) - 1))\mathbb{1}_{h_{2,1}^T\theta_{2,1}^*=0}) : a \in \mathbb{R}^{p_1}, \nu, \gamma \in \mathbb{R}^{p_1}, \max\{||a||, ||\nu||, ||\gamma||\} \le K\}$

3. $\omega_{13} : D_{p_1} \times D_{p_{21}} \times l^{\infty}(\mathcal{F}_{13}) \times \mathbb{R}^{p_{21}} \times \mathbb{R}^{p_{21}} \to \mathbb{R}$ is defined as

$$\begin{aligned} & \omega_{13}(\Sigma_1, \Sigma_{21,21}, \mu, \nu, \gamma) \\ := \; & 2\mu\Big[c^T\Sigma_1^{-1}B_1|H_{2,1}^T\gamma|\mathbb{1}(T(H_{2,1}, \nu+\gamma, \Sigma_{21,21}) > \chi)\times \\ & 1(H_{2,1}^T\gamma \times H_{2,1}^T(\nu+\gamma) < 0)\mathbb{1}_{H_{2,1}^T\theta_{2,1}^*=0}\Big] \end{aligned}$$

where $\mathcal{F}_{13} = \{f(b_1, h_{2,1}) = a^T b_1|h_{2,1}^T\gamma|\mathbb{1}(T(h_{2,1}, \nu+\gamma, \Sigma_{21,21}) > \chi)\mathbb{1}(h_{2,1}^T\gamma \times h_{2,1}^T(\nu+\gamma) < 0)\mathbb{1}_{h_{2,1}^T\theta_{2,1}^*=0}) : a \in \mathbb{R}^{p_1}, \nu, \gamma \in \mathbb{R}^{p_1}, \max\{||a||, ||\nu||, ||\gamma||\} \leq K\}$

4.$\rho_{11} : D_{p_1} \times D_{p_{21}} \times l^{\infty}(\tilde{\mathcal{F}}_{11}) \times \mathbb{R}^{p_{21}} \times \mathbb{R}^{p_{21}} \times \mathbb{R}^{p_{21}} \times \mathbb{R} \to \mathbb{R}$ is defined as

$$\begin{aligned} & \rho_{11}(\Sigma_1, \Sigma_{21,21}, \mu, \nu, \eta, \gamma, \lambda) \\ = \; & \mu\Big[c^T\Sigma_1^{-1}B_1(|H_{2,1}^T(\nu+\gamma)| - |H_{2,1}^T\gamma|)(2\mathbb{1}(T(H_{2,1}, \nu+\gamma, \Sigma_{21,21}) > \chi) - 1) \\ & \times(\mathbb{1}(T(H_{2,1}, \nu+\eta, \Sigma_{21,21}) \leq \lambda) - \mathbb{1}_{H_{2,1}^T\theta_{2,1}^*=0})\Big] \end{aligned}$$

where $\tilde{\mathcal{F}}_{11} = f(b_1, h_{2,1}) = a^T b_1(|h_{2,1}^T(\nu+\gamma)| - |h_{2,1}^T\gamma|)(2\mathbb{1}(T(h_{2,1}, \nu+\gamma, \Sigma_{21,21}) > \chi) - 1)(\mathbb{1}(T(h_{2,1}, \nu+\eta, \Sigma_{21,21}) \leq \lambda) - \mathbb{1}_{h_{2,1}^T\theta_{2,1}^*=0}) : a \in \mathbb{R}^{p_1}, \nu, \eta \in \mathbb{R}^{p_{21}}, \Sigma_{21,21} \in D_{p_{21}}, \max\{||a||, ||\nu||, ||\eta||, ||\gamma||\} \leq K\}$

5.$\rho_{12} : D_{p_1} \times D_{p_{21}} \times l^{\infty}(\tilde{\mathcal{F}}_{12}) \times \mathbb{R}^{p_{21}} \times \mathbb{R}^{p_{21}} \to \mathbb{R}$ is defined as

$$\begin{aligned} & \rho_{12}(\Sigma_1, \Sigma_{21,21}, \mu, \nu, \eta) \\ := \; & \mu(c^T\Sigma_1^{-1}B_1(|H_{2,1}^T(\nu+\eta) - |H_{2,1}^T\eta|)(2\mathbb{1}(T(H_{2,1}, \nu+\eta, \Sigma_{21,21}) > \chi) - 1) \\ & -H_{2,1}^T\nu)\mathbb{1}_{H_{2,1}^T\theta_{2,1}^*>0}\Big] \\ & +\mu(c^T\Sigma_1^{-1}B_1(|H_{2,1}^T(\nu+\eta) - |H_{2,1}^T\eta|))(2\mathbb{1}(T(H_{2,1}, \nu+\eta, \Sigma_{21,21}) > \chi) - 1) \\ & +H_{2,1}^T\nu)\mathbb{1}_{H_{2,1}^T\theta_{2,1}^*<0}\Big] \end{aligned}$$

where $\tilde{\mathcal{F}}_{12} = \{f(b_1, h_{2,1}) = a^T b_1((|h_{2,1}^T(\nu+\eta)| - |h_{2,1}^T\eta|)(2\mathbb{1}(T(h_{2,1}, \nu+\eta, \Sigma_{21,21}) > \chi) - 1) - h_{2,1}^T\nu)\mathbb{1}_{H_{2,1}^T\theta_{2,1}^* > 0} + a^T b_1((|h_{2,1}^T(\nu+\eta)| - |h_{2,1}^T\eta|)(2\mathbb{1}(T(h_{2,1}, \nu+\eta, \Sigma_{21,21}) > \chi) - 1) + h_{2,1}^T\nu)\mathbb{1}_{H_{2,1}^T\theta_{2,1}^* < 0} : a \in \mathbb{R}^{p_1}, \nu, \eta \in \mathbb{R}^{p_{21}}, \Sigma_{21,21} \in D_{p_{21}}, \max\{||a||, ||\nu||, ||\eta||\} \leq K\}$

6.$\rho_{13} : D_{p_1} \times D_{p_{21}} \times l^\infty(\tilde{\mathcal{F}}_{13}) \times \mathbb{R}^{p_{21}} \times \mathbb{R}^{p_{21}} \to \mathbb{R}$ is defined as

$$\begin{aligned} & \rho_{13}(\Sigma_1, \Sigma_{21,21}, \mu, \nu, \gamma) \\ := \; & 2\mu\Big[(c^T\Sigma_1^{-1}B_1(|H_{2,1}^T\gamma|\mathbb{1}(T(H_{2,1}, \nu+\gamma, \Sigma_{21,21}) > \chi)\times \\ & 1(H_{2,1}^T\gamma \times H_{2,1}^T(\nu+\gamma) < 0)\mathbb{1}_{H_{2,1}^T\theta_{2,1}^* \neq 0}\Big] \end{aligned}$$

where $\tilde{\mathcal{F}}_{13} = \{f(b_1, h_{2,1}) = a^T b_1(|h_{2,1}^T\gamma|\mathbb{1}(T(h_{2,1}, \nu+\gamma, \Sigma_{21,21}) > \chi)\mathbb{1}(h_{2,1}^T\gamma \times h_{2,1}^T(\nu+\gamma) < 0)\mathbb{1}_{h_{2,1}^T\theta_{2,1}^* \neq 0} : a \in \mathbb{R}^{p_1}, \nu, \gamma \in \mathbb{R}^{p_{21}}, \Sigma_{21,21} \in D_{p_{21}}, \max\{||a||, ||\nu||, ||\gamma||\} \leq K\}$

7.$\rho_{14} : D_{p_1} \times D_{p_{21}} \times l^\infty(\tilde{\mathcal{F}}_{14}) \times \mathbb{R}^{p_{21}} \times \mathbb{R}^{p_{21}} \times \mathbb{R}^{p_{21}} \times \mathbb{R} \to \mathbb{R}$ is defined as

$$\begin{aligned} & \rho_{14}(\Sigma_1, \Sigma_{21,21}, \mu, \nu, \eta, \gamma, \lambda) \\ := \; & 2\mu\Big[(c^T\Sigma_1^{-1}B_1(|H_{2,1}^T\gamma|\mathbb{1}(T(H_{2,1}, \nu+\gamma, \Sigma_{21,21}) > \chi)\mathbb{1}(H_{2,1}^T\gamma \times H_{2,1}^T(\nu+\gamma) < 0) \\ & \times(\mathbb{1}(T(H_{2,1}, \nu+\eta, \Sigma_{21,21}) \leq \lambda) - \mathbb{1}_{H_{2,1}^T\beta_{2,1}^* = 0})\Big] \end{aligned}$$

where $\tilde{\mathcal{F}}_{13} = \{f(b_1, h_{2,1}) = a^T b_1(|h_{2,1}^T\gamma|\mathbb{1}(T(h_{2,1}, \nu+\gamma, \Sigma_{21,21}) > \chi)\mathbb{1}(h_{2,1}^T\gamma \times h_{2,1}^T(\nu+\gamma) < 0)(\mathbb{1}(T(h_{2,1}, \nu+\eta, \Sigma_{21,21}) \leq \lambda) - \mathbb{1}_{h_{2,1}^T\beta_{2,1}^* = 0}) : a \in \mathbb{R}^{p_1}, \nu, \gamma, \eta \in \mathbb{R}^{p_{21}}, \Sigma_{21,21} \in D_{p_{21}}, \max\{||a||, ||\nu||, ||\gamma||, ||\eta||\} \leq K\}$

Using the following functions we have the following expressions for the first stage parameters:

$$
\begin{aligned}
c^T\sqrt{n}(\hat{\theta}_1-\hat{\theta}_1^*) &= \omega_{11}(\hat{\Sigma}_1,\hat{\Sigma}_{12},\hat{\Sigma}_{21,21},\mathbb{G}_n,\mathbb{P}_n,\sqrt{n}(\hat{\theta}_2-\theta_2^*),\sqrt{n}\theta_{2,1}^*,(\hat{\theta}_1^{*T},\theta_2^{*T})^T)\\
&\quad+\omega_{12}(\hat{\Sigma}_1,\hat{\Sigma}_{21,21},\mathbb{P}_n,\sqrt{n}(\hat{\theta}_{2,1}-\theta_{2,1}^*),\sqrt{n}\theta_{2,1}^*)\\
&\quad+\omega_{13}(\hat{\Sigma}_1,\hat{\Sigma}_{21,21},\mathbb{P}_n,\sqrt{n}(\hat{\theta}_{2,1}-\theta_{2,1}^*),\sqrt{n}\theta_{2,1}^*)\\
&\quad+\rho_{12}(\hat{\Sigma}_1,\hat{\Sigma}_{21,21},\mathbb{P}_n,\sqrt{n}(\hat{\theta}_{2,1}-\theta_{2,1}^*),\sqrt{n}\theta_{2,1}^*)\\
&\quad+\rho_{13}(\hat{\Sigma}_1,\hat{\Sigma}_{21,21},\mathbb{P}_n,\sqrt{n}(\hat{\theta}_{2,1}-\theta_{2,1}^*),\sqrt{n}\theta_{2,1}^*)
\end{aligned}
$$

$$
\begin{aligned}
c^T\sqrt{n}(\hat{\theta}_1-\hat{\theta}_{1,n}^*) &= \omega_{11}(\hat{\Sigma}_1,\hat{\Sigma}_{12},\sqrt{n}(\mathbb{P}_n-P_n),\mathbb{P}_n,\sqrt{n}(\hat{\theta}_2-\theta_{2,n}^*),(\theta_{1,n}^{*T},\theta_{2,n}^{*T})^T)\\
&\quad+\omega_{12}(\hat{\Sigma}_1,\hat{\Sigma}_{21,21},\mathbb{P}_n,\sqrt{n}(\hat{\theta}_{2,1}-\theta_{2,1,n}^*),\sqrt{n}\theta_{2,1,n}^*)\\
&\quad+\omega_{13}(\hat{\Sigma}_1,\hat{\Sigma}_{21,21},\mathbb{P}_n,\sqrt{n}(\hat{\theta}_{2,1}-\theta_{2,1,n}^*),\sqrt{n}\theta_{2,1,n}^*)\\
&\quad+\rho_{12}(\hat{\Sigma}_1,\hat{\Sigma}_{21,21},\mathbb{P}_n,\sqrt{n}(\hat{\theta}_{2,1}-\theta_{2,1,n}^*),\sqrt{n}\theta_{2,1,n}^*)\\
&\quad+\rho_{13}(\hat{\Sigma}_1,\hat{\Sigma}_{21,21},\mathbb{P}_n,\sqrt{n}(\hat{\theta}_{2,1}-\theta_{2,1,n}^*),\sqrt{n}\theta_{2,1,n}^*)
\end{aligned}
$$

Similarly we can express the upper bound $\mathcal{U}(c)$ in terms of the above functions:

$$
\begin{aligned}
\mathcal{U}(c) &= \omega_{11}(\hat{\Sigma}_1,\hat{\Sigma}_{12},\hat{\Sigma}_{21,21},\mathbb{G}_n,\mathbb{P}_n,\sqrt{n}(\hat{\theta}_2-\theta_2^*),\sqrt{n}\theta_{2,1}^*,(\hat{\theta}_1^{*T},\theta_2^{*T})^T)\\
&\quad-\rho_{11}(\hat{\Sigma}_1,\hat{\Sigma}_{21,21},\mathbb{P}_n,\sqrt{n}(\hat{\theta}_{2,1}-\theta_{2,1}^*),\sqrt{n}\theta_{2,1}^*,\sqrt{n}\theta_{2,1}^*,\lambda_n)\\
&\quad+\rho_{12}(\hat{\Sigma}_1,\mathbb{P}_n,\sqrt{n}(\hat{\theta}_{2,1}-\theta_{2,1}^*),\sqrt{n}\theta_{2,1}^*)\\
&\quad+\rho_{13}(\hat{\Sigma}_1,\hat{\Sigma}_{21,21},\mathbb{P}_n,\sqrt{n}(\hat{\theta}_{2,1}-\theta_{2,1}^*),\sqrt{n}\theta_{2,1}^*)\\
&\quad+\rho_{14}(\hat{\Sigma}_1,\hat{\Sigma}_{21,21},\mathbb{P}_n,\sqrt{n}(\hat{\theta}_{2,1}-\theta_{2,1}^*),\sqrt{n}\theta_{2,1}^*,\sqrt{n}\theta_{2,1}^*,\lambda_n)\\
&\quad+\sup_{\gamma\in\mathbb{R}^{dim(\theta_{2,1}^*)}}\Big[\omega_{12}(\hat{\Sigma}_1,\hat{\Sigma}_{21,21},\mathbb{P}_n,\sqrt{n}(\hat{\theta}_{2,1}-\theta_{2,1}^*),\gamma)\\
&\quad\omega_{13}(\hat{\Sigma}_1,\hat{\Sigma}_{21,21},\mathbb{P}_n,\sqrt{n}(\hat{\theta}_{2,1}-\theta_{2,1}^*),\gamma)\\
&\quad+\rho_{11}(\hat{\Sigma}_1,\hat{\Sigma}_{21,21},\mathbb{P}_n,\sqrt{n}(\hat{\theta}_{2,1}-\theta_{2,1}^*),\sqrt{n}\theta_{2,1}^*,\gamma,\lambda_n)\\
&\quad+\rho_{14}(\hat{\Sigma}_1,\hat{\Sigma}_{21,21},\mathbb{P}_n,\sqrt{n}(\hat{\theta}_{2,1}-\theta_{2,1}^*),\sqrt{n}\theta_{2,1}^*,\gamma,\lambda_n)\Big]
\end{aligned}
$$

We will also make use of the following alternative expression for the upper bound $\mathcal{U}(c)$ under P_n:

$$\begin{aligned}
\mathcal{U}(c) &= \omega_{11}(\hat{\Sigma}_1, \hat{\Sigma}_{12}, \sqrt{n}(\mathbb{P}_n - P_n), \mathbb{P}_n, \sqrt{n}(\hat{\theta}_2 - \theta^*_{2,n}), \sqrt{n}\theta^*_{2,1,n}, (\theta^{*T}_{1,n}, \theta^{*T}_{2,n})^T) \\
&\quad -\rho_{10}(\hat{\Sigma}_1, \hat{\Sigma}_{21,21}, \mathbb{P}_n, \sqrt{n}(\hat{\theta}_{2,1} - \theta^*_{2,1,n}), \sqrt{n}\theta^*_{2,1,n}, \sqrt{n}\theta^*_{2,1,n}, \lambda_n) \\
&\quad +\rho_{12}(\hat{\Sigma}_1, \mathbb{P}_n, \sqrt{n}(\hat{\theta}_{2,1} - \theta^*_{2,1,n}), \sqrt{n}\theta^*_{2,1,n}) \\
&\quad +\rho_{13}(\hat{\Sigma}_1, \hat{\Sigma}_{21,21}, \mathbb{P}_n, \sqrt{n}(\hat{\theta}_{2,1} - \theta^*_{2,1,n}), \sqrt{n}\theta^*_{2,1,n}) \\
&\quad +\rho_{14}(\hat{\Sigma}_1, \hat{\Sigma}_{21,21}, \mathbb{P}_n, \sqrt{n}(\hat{\theta}_{2,1} - \theta^*_{2,1,n}), \sqrt{n}\theta^*_{2,1,n}, \sqrt{n}\theta^*_{2,1,n}, \lambda_n) \\
&\quad + \sup_{\gamma \in \mathbb{R}^{dim(\theta^*_{2,1})}} \Big[\omega_{12}(\hat{\Sigma}_1, \hat{\Sigma}_{21,21}, \mathbb{P}_n, \sqrt{n}(\hat{\theta}_{2,1} - \theta^*_{2,1,n}), \gamma) \\
&\quad \omega_{13}(\hat{\Sigma}_1, \hat{\Sigma}_{21,21}, \mathbb{P}_n, \sqrt{n}(\hat{\theta}_{2,1} - \theta^*_{2,1,n}), \gamma) \\
&\quad +\rho_{11}(\hat{\Sigma}_1, \hat{\Sigma}_{21,21}, \mathbb{P}_n, \sqrt{n}(\hat{\theta}_{2,1} - \theta^*_{2,1,n}), \sqrt{n}\theta^*_{2,1,n}, \gamma, \lambda_n) \\
&\quad +\rho_{14}(\hat{\Sigma}_1, \hat{\Sigma}_{21,21}, \mathbb{P}_n, \sqrt{n}(\hat{\theta}_{2,1} - \theta^*_{2,1,n}), \sqrt{n}\theta^*_{2,1,n}, \gamma, \lambda_n)\Big]
\end{aligned}$$

Similarly we will make use of following expression for the bootstrap analog of the upper bound

$$\begin{aligned}
\mathcal{U}^{(b)}(c) &= \omega_{11}(\hat{\Sigma}^{(b)}_1, \hat{\Sigma}^{(b)}_{12}, \sqrt{n}(\mathbb{P}^{(b)}_n - \mathbb{P}_n), \mathbb{P}^{(b)}_n, \sqrt{n}(\hat{\theta}^{(b)}_2 - \hat{\theta}_2), \sqrt{n}\hat{\theta}_{2,1}, (\hat{\theta}^T_1, \hat{\theta}^T_2)^T) \\
&\quad -\rho_{10}(\hat{\Sigma}^{(b)}_1, \hat{\Sigma}^{(b)}_{21,21}, \mathbb{P}^{(b)}_n, \sqrt{n}(\hat{\theta}^{(b)}_{2,1} - \hat{\theta}_{2,1}), \sqrt{n}\hat{\theta}_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \lambda_n) \\
&\quad +\rho_{12}(\hat{\Sigma}^{(b)}_1, \mathbb{P}^{(b)}_n, \sqrt{n}(\hat{\theta}^{(b)}_{2,1} - \hat{\theta}_{2,1}), \sqrt{n}\hat{\theta}_{2,1}) \\
&\quad +\rho_{13}(\hat{\Sigma}^{(b)}_1, \hat{\Sigma}^{(b)}_{21,21}, \mathbb{P}^{(b)}_n, \sqrt{n}(\hat{\theta}^{(b)}_{2,1} - \hat{\theta}_{2,1}), \sqrt{n}\hat{\theta}_{2,1}) \\
&\quad +\rho_{14}(\hat{\Sigma}^{(b)}_1, \hat{\Sigma}^{(b)}_{21,21}, \mathbb{P}^{(b)}_n, \sqrt{n}(\hat{\theta}^{(b)}_{2,1} - \hat{\theta}_{2,1}), \sqrt{n}\hat{\theta}_{2,1}, \sqrt{n}\hat{\theta}_{2,1}, \lambda_n) \\
&\quad + \sup_{\gamma \in \mathbb{R}^{dim(\theta^*_{2,1})}} \Big[\omega_{12}(\hat{\Sigma}^{(b)}_1, \hat{\Sigma}^{(b)}_{21,21}, \mathbb{P}^{(b)}_n, \sqrt{n}(\hat{\theta}^{(b)}_{2,1} - \hat{\theta}_{2,1}), \gamma) \\
&\quad \omega_{13}(\hat{\Sigma}^{(b)}_1, \hat{\Sigma}^{(b)}_{21,21}, \mathbb{P}^{(b)}_n, \sqrt{n}(\hat{\theta}^{(b)}_{2,1} - \hat{\theta}_{2,1}), \gamma) \\
&\quad +\rho_{11}(\hat{\Sigma}^{(b)}_1, \hat{\Sigma}^{(b)}_{21,21}, \mathbb{P}^{(b)}_n, \sqrt{n}(\hat{\theta}^{(b)}_{2,1} - \hat{\theta}_{2,1}), \sqrt{n}\hat{\theta}_{2,1}, \gamma, \lambda_n) \\
&\quad +\rho_{14}(\hat{\Sigma}^{(b)}_1, \hat{\Sigma}^{(b)}_{21,21}, \mathbb{P}^{(b)}_n, \sqrt{n}(\hat{\theta}^{(b)}_{2,1} - \hat{\theta}_{2,1}), \sqrt{n}\hat{\theta}_{2,1}, \gamma, \lambda_n)\Big]
\end{aligned}$$

Below we argue that ρ_{11} through ρ_{14} are negligible and ω_{11} through ω_{13} are continuous in such a fashion so as to facilitate the use of continuous mapping theorem.

First we show the negligibility of the ρs. The function ρ_{11} is the most difficult to handle so we address it here and omit the proof of ρ_{12}, ρ_{13} and ρ_{14}.

Theorem A.3. *Assume (A1)-(A4). Then*

1. $\sup_{\gamma\in\mathbb{R}^{p_{21}}} |\rho_{11}(\hat{\Sigma}_1, \hat{\Sigma}_{21,21}, \mathbb{P}_n, \sqrt{n}(\hat{\theta}_{2,1} - \theta^*_{2,1}), \sqrt{n}\theta^*_{2,1}, \gamma, \lambda_n)| \to_P 0.$

2. $\sup_{\gamma\in\mathbb{R}^{p_{21}}} |\rho_{11}(\hat{\Sigma}^{(b)}_1, \hat{\Sigma}^{(b)}_{21,21}, \hat{\mathbb{P}}^{(b)}_n, \sqrt{n}(\hat{\theta}^{(b)}_{2,1} - \hat{\theta}_{2,1}), \sqrt{n}\hat{\theta}_{2,1}, \gamma, \lambda_n)| \to_{P_M} 0$ *almost surely* P.

3. $\sup_{\gamma\in\mathbb{R}^{p_{21}}} |\rho_{11}(\hat{\Sigma}_1, \hat{\Sigma}_{21,21}, \mathbb{P}_n, \sqrt{n}(\hat{\theta}_{2,1} - \theta^*_{2,1,n}), \sqrt{n}\theta^*_{2,1,n}, \gamma, \lambda_n)| \to_{P_n} 0,$

Proof. First, it's easy to show that $||H^T_{2,1}\nu - H^T_{2,1}\gamma| - |H^T_{2,1}\gamma|| \leq |H^T_{2,1}\nu|$. Thus for any probability measure μ in $l^\infty(\tilde{\mathcal{F}}_{11})$,

$$
\begin{aligned}
& |\rho_{11}(\Sigma_1, \Sigma_{21,21}, \mu, \nu, \eta, \gamma, \lambda)| \\
\leq\ & K\Big\{\mu(||B_1||\ ||H_{2,1}||\mathbb{1}_{H^*_{2,1}\theta^*_{2,1}=0, H^T_{2,1}\eta/||H_{2,1}||>\sqrt{\lambda k}-K}) \\
& +\mu(||B_1||\ ||H_{2,1}||\mathbb{1}_{H^*_{2,1}\theta^*_{2,1}=0, H^T_{2,1}\eta/||H_{2,1}||<\sqrt{\lambda k}-K}) \\
& +\mu(||B_1||\ ||H_{2,1}||\mathbb{1}_{H^*_{2,1}\theta^*_{2,1}\neq 0, \sqrt{\lambda k}-K\leq H^T_{2,1}\eta/||H_{2,1}||\leq\sqrt{\lambda k}+K})\Big\}
\end{aligned}
$$

for a sufficiently large constant $K > 0$ and a sufficiently small constant $k \in (0, 1)$. The rest of the proof is exactly the same as those in *Laber et al.* (2010). □

Next we would like to show that ω_{11} is continuous at points in
$(\Sigma_{1,\infty}, \Sigma_{12,\infty}, C_b(\mathcal{F}_{11}), P, \mathbb{R}^{p_2}, (\hat{\theta}^{*T}_1, \theta^{*T}_2)^T)$, $\omega_{12}(\cdot,\cdot,\cdot,\cdot,\sqrt{n}\theta^*_{2,1})$,$\omega_{12}(\cdot,\cdot,\cdot,\cdot,\sqrt{n}\theta^*_{2,1,n})$,
$\omega_{13}(\cdot,\cdot,\cdot,\cdot,\sqrt{n}\theta^*_{2,1})$ and $\omega_{13}(\cdot,\cdot,\cdot,\cdot,\sqrt{n}\theta^*_{2,1,n})$ are continuous at points in
$(\Sigma_{1,\infty}, \Sigma_{21,21,\infty}, P, \mathbb{R}^{p_{21}})$, and $\omega'_{12} := \sup_{\gamma\in\mathbb{R}^{p_{21}}} \omega_{12}(\Sigma_1, \Sigma_{21,21}, \mu, \nu, \gamma)$,
$\omega'_{13} := \sup_{\gamma\in\mathbb{R}^{p_{21}}} \omega_{13}(\Sigma_1, \Sigma_{21,21}, \mu, \nu, \gamma)$ are continuous at points in $(\Sigma_{1,\infty}, \Sigma_{21,21,\infty}, P, \mathbb{R}^{p_{21}})$.

To prove the desired continuity of ω_{12} and ω'_{12}, we will establish the stronger result that ω_{13} is continuous at points $(\Sigma_{1,\infty}, \Sigma_{21,21,\infty}, P, \mathbb{R}^{p_{21}}, \gamma, +\infty)$ uniformly in γ. That is, for any $\Sigma_n \to \Sigma_{1,\infty}, \Sigma_{21,21,n} \to \Sigma_{21,21,\infty}$, probability measure $\mu_n \to P$ and $\nu_n \to \nu$ (where ν equals one of the following: $\sqrt{n}(\hat{\theta}_{2,1} - \theta^*_{2,1}), \sqrt{n}(\hat{\theta}_{2,1} - \theta^*_{2,1,n}), \sqrt{n}(\hat{\theta}^{(b)}_{2,1} - \hat{\theta}_{2,1})$), we have

Theorem A.4. $\sup_\gamma |\omega_{12}(\Sigma_n, \Sigma_{21,21,n}, \mu_n, \nu_n, \gamma) - \omega_{12}(\Sigma_{1,\infty}, \Sigma_{21,21,\infty}, P, \nu, \gamma)| \to 0$

Proof.

$$
\begin{aligned}
& |\omega_{12}(\Sigma_n, \Sigma_{21,21,n}, \mu_n, \nu_n, \gamma) - \omega_{12}(\Sigma_{1,\infty}, \Sigma_{21,21,\infty}, P, \nu, \gamma)| \\
\leq\ & |\omega_{12}(\Sigma_n, \Sigma_{21,21,n}, \mu_n, \nu_n, \gamma) - \omega_{12}(\Sigma_n, \Sigma_{21,21,n}, \mu_n, \nu, \gamma)| \\
& + |\omega_{12}(\Sigma_n, \Sigma_{21,21,n}, \mu_n, \nu, \gamma) - \omega_{12}(\Sigma_n, \Sigma_{21,21,n}, P, \nu, \gamma)| \\
& + |\omega_{12}(\Sigma_n, \Sigma_{21,21,n}, P, \nu, \gamma) - \omega_{12}(\Sigma_n, \Sigma_{21,21,\infty}, P, \nu, \gamma)| \\
& + |\omega_{12}(\Sigma_n, \Sigma_{21,21,\infty}, P, \nu, \gamma) - \omega_{12}(\Sigma_{1,\infty}, \Sigma_{21,21,\infty}, P, \nu, \gamma)| \\
=\ & I + II + III + IV
\end{aligned}
$$

$$
\begin{aligned}
& I \\
= \; & \mu_n \Big[c^T \Sigma_n^{-1} B_1 (|H_{2,1}^T(\nu_n + \gamma)| - |H_{2,1}^T \gamma|)(2\mathbb{1}(T(H_{2,1}, \nu_n + \gamma, \Sigma_{21,21,n}) > \chi) - 1) \times \\
& \mathbb{1}_{H_{2,1}^T \theta_{2,1}^* = 0} \Big] \\
& - \mu_n \Big[c^T \Sigma_n^{-1} B_1 (|H_{2,1}^T(\nu + \gamma)| - |H_{2,1}^T \gamma|)(2\mathbb{1}(T(H_{2,1}, \nu + \gamma, \Sigma_{21,21,n}) > \chi) - 1) \times \\
& \mathbb{1}_{H_{2,1}^T \theta_{2,1}^* = 0} \Big] \\
= \; & \mu_n \Big[c^T \Sigma_n^{-1} B_1 (|H_{2,1}^T(\nu_n + \gamma)| - |H_{2,1}^T(\nu + \gamma)|)(2\mathbb{1}(T(H_{2,1}, \nu_n + \gamma, \Sigma_{21,21,n}) > \chi) \\
& - 1) \mathbb{1}_{H_{2,1}^T \theta_{2,1}^* = 0} \Big] \\
& + 2\mu_n \Big[c^T \Sigma_n^{-1} B_1 (|H_{2,1}^T(\nu + \gamma)| - |H_{2,1}^T \gamma|) \\
& \times (\mathbb{1}(T(H_{2,1}, \nu_n + \gamma, \Sigma_{21,21,n}) > \chi) - \mathbb{1}(T(H_{2,1}, \nu + \gamma, \Sigma_{21,21,n}) > \chi)) \mathbb{1}_{H_{2,1}^T \theta_{2,1}^* = 0} \Big] \\
= \; & I_1 + I_2
\end{aligned}
$$

We have $I_1 \leq \mu_n(|c^T \Sigma_n^{-1} B_1 |H_{2,1}^T(\nu_n - \nu)||) \leq ||c||\ ||\Sigma_n^{-1}|| \mu_n(||B_1||\ ||H_{2,1}||) ||\nu_n - \nu|| = o(1)$ since we have $||\Sigma_n^{-1}||$ is bounded above for sufficiently large n.

Let $M := ||c||\ ||\Sigma_n^{-1}||\ ||B_1||\ ||H_{2,1}||\ ||\nu||$, then

$$
\begin{aligned}
& I_2 \\
\leq\ & ||c||\ ||\Sigma_n^{-1}||\mu_n\Big[||B_1||\ ||H_{2,1}||\ ||\nu||(\mathbb{1}(T(H_{2,1},\nu_n+\gamma,\Sigma_{21,21,n})>\chi) \\
& -\mathbb{1}(T(H_{2,1},\nu+\gamma,\Sigma_{21,21,n})>\chi))\Big] \\
=\ & \mu_n\Big[M|\mathbb{1}(T(H_{2,1},\nu_n+\gamma,\Sigma_{21,21,n})>\chi)-\mathbb{1}(T(H_{2,1},\nu+\gamma,\Sigma_{21,21,n})>\chi)|\Big] \\
=\ & \mu_n\Big[M|\mathbb{1}(T(H_{2,1},\nu_n+\gamma,\Sigma_{21,21,n})>\chi)-\mathbb{1}(T(H_{2,1},\nu+\gamma,\Sigma_{21,21,n})>\chi)| \\
& \times\mathbb{1}(T(H_{2,1},\nu+\gamma,\Sigma_{21,21,\infty})>\chi+\delta)\Big] \\
=\ & \mu_n\Big[M|\mathbb{1}(T(H_{2,1},\nu_n+\gamma,\Sigma_{21,21,n})>\chi)-\mathbb{1}(T(H_{2,1},\nu+\gamma,\Sigma_{21,21,n})>\chi)| \\
& \times\mathbb{1}(T(H_{2,1},\nu+\gamma,\Sigma_{21,21,\infty})<\chi-\delta)\Big] \\
=\ & \mu_n\Big[M|\mathbb{1}(T(H_{2,1},\nu_n+\gamma,\Sigma_{21,21,n})>\chi)-\mathbb{1}(T(H_{2,1},\nu+\gamma,\Sigma_{21,21,n})>\chi)| \\
& \times\mathbb{1}(\chi-\delta\leq T(H_{2,1},\nu+\gamma,\Sigma_{21,21,\infty})\leq\chi+\delta)\Big] \\
:=\ & I_{21}(\delta)+I_{22}(\delta)+I_{23}(\delta)
\end{aligned}
$$

For any positive number δ.

$I_{23}(\delta) \leq 2M\mu_n\mathbb{1}(\chi-\delta \leq T(H_{2,1},\nu+\gamma,\Sigma_{21,21,\infty}) \leq \chi+\delta) \leq 2M||\mu_n-P||+2MP\mathbb{1}(\chi-\delta \leq T(H_{2,1},\nu+\gamma,\Sigma_{21,21,\infty}) \leq \chi+\delta)$. When $\delta \to 0$, the second term converges to $2MP\mathbb{1}((T(H_{2,1},\nu+\gamma,\Sigma_{21,21,\infty}) = \chi+\delta) = 0$ since we have $m\mathbb{1}((T(H_{2,1},\nu+\gamma,\Sigma_{21,21,\infty}) = \chi+\delta) = 0$ and we have the absolute continuity assumption. The first term goes to zero as n increases. So for any $\epsilon > 0$, there exists N_1 sufficiently large and $\delta > 0$ sufficiently small such that $I_{23}(\delta) < \epsilon/3$ for all $n > N_1$.

For this fixed δ, as $\nu_n \to \nu$ and $\Sigma_{21,21,n} \to \Sigma_{21,21,\infty}$ and T is continuous with respect to all its components we have both $T(H_{2,1},\nu_n+\gamma,\Sigma_{21,21,n})$ and $T(H_{2,1},\nu+\gamma,\Sigma_{21,21,n})$ converge to $T(H_{2,1},\nu+\gamma,\Sigma_{21,21,\infty})$ and the speed is independent of γ. So there exists N_2 such that $P\Big(|T(H_{2,1},\nu_n+\gamma,\Sigma_{21,21,n})-T(H_{2,1},\nu+\gamma,\Sigma_{21,21,\infty})| < \delta/2\Big) \geq 1-\epsilon/9$ and $P\Big(|T(H_{2,1},\nu+\gamma,\Sigma_{21,21,n})-T(H_{2,1},\nu+\gamma,\Sigma_{21,21,\infty})| < \delta/2\Big) \geq 1-$

$\epsilon/9$ for all $n > N_2$. In such case, $I_{23}(\delta) \leq 2M||\mu_n - P|| + MP\Big[|\mathbb{1}(T(H_{2,1}, \nu_n + \gamma, \Sigma_{21,21,n}) > \chi) - \mathbb{1}(T(H_{2,1}, \nu + \gamma, \Sigma_{21,21,n}) > \chi)|\mathbb{1}(T(H_{2,1}, \nu + \gamma, \Sigma_{21,21,\infty}) > \chi + \delta)\Big]$. The first term is no larger than $\epsilon/9$ when $n > N_3$ for some large N_3, the second term is no larger than $MP\Big[\mathbb{1}(T(H_{2,1}, \nu_n + \gamma, \Sigma_{21,21,n}) \leq \chi)\mathbb{1}(T(H_{2,1}, \nu + \gamma, \Sigma_{21,21,\infty}) > \chi + \delta)\Big] + MP\Big[\mathbb{1}(T(H_{2,1}, \nu + \gamma, \Sigma_{21,21,n}) \leq \chi)\mathbb{1}(T(H_{2,1}, \nu + \gamma, \Sigma_{21,21,\infty}) > \chi + \delta)\Big] = MP\Big(|T(H_{2,1}, \nu_n + \gamma, \Sigma_{21,21,n}) - T(H_{2,1}, \nu + \gamma, \Sigma_{21,21,\infty})| < \delta/2\Big) + MP\Big(|T(H_{2,1}, \nu + \gamma, \Sigma_{21,21,n}) - T(H_{2,1}, \nu + \gamma, \Sigma_{21,21,\infty})| < \delta/2\Big) \leq \epsilon/9 + \epsilon/9$. So we have $I_{21}(\delta) \leq \epsilon/3$ for all $n > \max(N_2, N_3)$. Similarly we have $I_{22}(\delta) \leq \epsilon/3$ for all $n > \max(N_2, N_3)$.
So for all $n > \max(N_1, N_2, N_3)$ we have that $I_2 < \epsilon$. So we have $I \to 0$.

$$
\begin{aligned}
& II \\
= \; & |(\mu_n - P)((c^T\Sigma_n^{-1}B_1(|H_{2,1}^T\nu + H_{2,1}^T\gamma| - \\
& |H_{2,1}^T\gamma|)(2\mathbb{1}(T(H_{2,1}, \nu + \gamma, \Sigma_{21,21,n}) > \chi) - 1)\mathbb{1}_{H_{2,1}^T\theta_{2,1}^* = 0} \\
\leq \; & |(\mu - P)|c^T\Sigma_n^{-1}B_1(|H_{2,1}^T\nu + H_{2,1}^T\gamma| - |H_{2,1}^T\gamma|)||
\end{aligned}
$$

If $||\nu|| = 0$ obviously $II = 0$. Otherwise, $II \leq |(\mu_n - P)|((c^T\Sigma_n^{-1}B_1(|H_{2,1}^T\nu + H_{2,1}^T\gamma|/||\nu|| - |H_{2,1}^T\gamma|/||\nu||)| \cdot ||\nu|| \leq ||\mu_n - P|| \cdot ||\nu|| \to 0$.

$$
\begin{aligned}
& III \\
= \; & 2P(c^T\Sigma_n^{-1}B_1(|H_{2,1}^T(\nu + \gamma)| - |H_{2,1}^T\gamma|) \\
& \times\big(\mathbb{1}(T(H_{2,1}, \nu + \gamma, \Sigma_{21,21,n}) > \chi) - \mathbb{1}(T(H_{2,1}, \nu + \gamma, \Sigma_{21,21,\infty}) > \chi)\big)\mathbb{1}_{H_{2,1}^T\theta_{2,1}^* = 0}
\end{aligned}
$$

Following similar proof for the convergence of I_2 we have $III \to 0$ uniformly over γ as $n \to \infty$.
$IV \leq P\big(||\Sigma_n^{-1} - \Sigma_{1,\infty}^{-1}||(||c^T||\ ||B_1||\ ||H_{2,1}^T||\ ||\nu||)\big) \to 0$.

This established the continuity of ω_{12} and hence ω'_{12}. The continuity of ω_{11} and ω_{13} can be established through similar arguments and is therefore omitted. □

A.2 The proof of lemma III.8

First we prove the lemma for the case where the treatment can take two values.

Lemma A.5. *Suppose the treatment A can take two values 1 and -1. We are using the linear model as*

$$E(Y) = h_0\beta_0 + h_1\beta_1 \times f(A) \tag{A.1}$$

where $h_1 \subset h_0$ and f is any function satisfying $f(1) \neq f(-1)$. Denote the OLS solution of (A.1) as

$$\hat{\beta}^{(f)} = (\hat{\beta}_0^{(f)}, \hat{\beta}_1^{(f)})$$

Then for any new observation $c = (c_0, c_1 f(A))$, we have that both $c_0\hat{\beta}_0^{(f)} + c_1\hat{\beta}_1^{(f)} \times f(1)$ and $c_0\hat{\beta}_0^{(f)} + c_1\hat{\beta}_1^{(f)} \times f(-1)$ do not depend on f.

Proof. Let our observation have the form H_0, H_1 and Y. The proof of this lemma is equivalent to prove that for every choice of f, the two terms above equal to the case when f is identity function.

First without of loss we assume that the first n_1 samples received $A = 1$ and the last n_2 samples received $A = -1$. (Thus $n = n_1 + n_2$) Denote p_1 as the dimension of β_1. As we have $H_1 \subset H_0$, without loss we assume that the last p_1 columns of H_0 is H_1 and rewrite $H_0 = (H_{0,0}, H_1)$. Let X_I and X_f denote the design matrix when f equals to identity function and general f, i.e.

$$X_I = \begin{pmatrix} H_{0,0,1} & H_{1,1} & H_{1,1} \\ H_{0,0,-1} & H_{1,-1} & -H_{1,-1} \end{pmatrix} \quad X_f = \begin{pmatrix} H_{0,0,1} & H_{1,1} & H_{1,1}f(1) \\ H_{0,0,-1} & H_{1,-1} & H_{1,-1}f(-1) \end{pmatrix}$$

where the last subscript of Hs means the treatment they receive. We denote the

number of columns of $H_{0,0,1}$ as p_0. So the dimension of β is $p = p_0 + 2p_1$. It's not hard to show that we have $X_f = X_I \times Q$ where Q is a p by p matrix with the form

$$Q = \begin{pmatrix} I_{p_0 \times p_0} & O & O \\ O & I_{p_1 \times p_1} & J_1 \\ O & O & J_2 \end{pmatrix} J_1 = \frac{f(1) + f(-1)}{2} I_{p_1 \times p_1}, J_2 = \frac{f(1) - f(-1)}{2} I_{p_1 \times p_1},$$

Another thing to mention is Q is invertible.

Denote the estimator of β using f being identity matrix as $\hat{\beta}^{(I)}$, we will only show that $c_0\hat{\beta}_0^{(f)} + c_1\hat{\beta}_1^{(f)} \times f(1) = c_0\hat{\beta}_0^{(I)} + c_1\hat{\beta}_1^{(I)}$. The other equality can be proved using exactly the same argument.

Denote $c_I = (c_0, c_1)$ and $c_f = (c_0, c_1 f(1))$, we also have $c_f = c_I Q$. Then

$$\begin{aligned} & c_0\hat{\beta}_0^{(f)} + c_1\hat{\beta}_1^{(f)} \times f(1) = c_f\hat{\beta}^{(f)} \\ = \; & c_f(X_f^T X_f)^{-1} X_f^T Y \\ = \; & c_I Q(Q^T X_I^T X_I Q)^{-1} Q^T X_I Y \\ = \; & c_I Q Q^{-1}(X_I^T X_I)^{-1} Q^{-T} Q^T X_I^T Y \\ = \; & c_I(X_I^T X_I)^{-1} X_I^T Y \\ = \; & c_I\hat{\beta}^{(I)} = c_0\hat{\beta}_0^{(I)} + c_1\hat{\beta}_1^{(I)} \end{aligned}$$

And And

$$\begin{aligned} & Var(c_f\hat{\beta}^{(f)}) \\ = \; & c_f(X_f^T X_f)^{-1} c_f^T \times (Y - X_f\hat{\beta}^{(f)})^T (Y - X_f\hat{\beta}^{(f)})/n \\ = \; & c_I Q(Q^T X_I^T X_I Q)^{-1} Q^T c_I^T \times (Y - X_I\hat{\beta}^{(I)})^T (Y - X_I\hat{\beta}^{(I)})/n \\ = \; & c_I(X_I^T X_I)^{-1} c_I^T \times (Y - X_I\hat{\beta}^{(I)})^T (Y - X_I\hat{\beta}^{(I)})/n \\ = \; & Var(c_I\hat{\beta}^{(I)}) \end{aligned}$$

□

Now we prove lemma A.6

Similarly, we first make a special choice of f_1 and f_2. Here we use $f_1(A) = \mathbb{1}(A = 1)$ and $f_2(A) = \mathbb{1}(A = 2)$. Denote the corresponding design matrix as X_0, again without loss we assume that the number of columns of $H_{2,1} = H_{2,2}$ as p_1 and the last p_1 columns of $H_{2,0}$ are $H_{2,1}$. i.e., $H_{2,0} = (H_{2,0,0}, H_{2,1})$. Denote the number of columns of $H_{2,0,0}$ as p_0. So the number of columns of H_2 is $p = p_0 + 3p_1$. Again without loss we assume that the first n_0 rows of observations receive $A_2 = 0$, the next n_1 rows receive $A_2 = 1$ and the last n_3 rows receive $A_2 = 2$. $n = n_1 + n_2 + n_3$. We can rewrite X_0 as

$$X_0 = \begin{pmatrix} H_{2,0,0,0} & H_{2,1,0} & O & O \\ H_{2,0,0,1} & H_{2,1,1} & H_{2,1,1} & O \\ H_{2,0,0,2} & H_{2,1,2} & O & H_{2,1,2} \end{pmatrix}$$

where $H_{2,0,0,i}$ is n_i by p_0 matrix meaning the rows of $H_{2,0,0}$ receiving $A_2 = i$. Let X_f be the design matrix when we use the general f_1 and f_2. Again all we need to prove is that the estimated effect and variance will be equal to the case when we choose $f_1(A) = \mathbb{1}(A = 1)$ and $f_2(A) = \mathbb{1}(A = 2)$. We have

$$X_f = \begin{pmatrix} H_{2,0,0,0} & H_{2,1,0} & H_{2,1,0}f_1(0) & H_{2,1,0}f_2(0) \\ H_{2,0,0,1} & H_{2,1,1} & H_{2,1,1}f_1(1) & H_{2,1,1}f_2(1) \\ H_{2,0,0,2} & H_{2,1,2} & H_{2,1,2}f_1(2) & H_{2,1,2}f_2(2) \end{pmatrix}$$

We have $X_f = X_0 Q$ where

$$Q = \begin{pmatrix} I_{p_0 \times p_0} & O & O & O \\ O & I_{p_1 \times p_1} & If_1(0) & If_2(0) \\ O & -I & I_{p_1 \times p_1}(f_1(1) - f_1(0)) & I(f_2(1) - f_2(0)) \\ O & -I & I(f_1(2) - f_1(0)) & I_{p_1 \times p_1}(f_2(2) - f_2(0)) \end{pmatrix}$$

Again Q is invertible. For any new sample with c_0 under the coding of which X_0 is the design matrix, it becomes $c_f = c_0 Q$ under the coding of which X_f is the design matrix. So we have

$$\begin{aligned} & c_f \hat{\beta}^{(f)} \\ = \; & c_f (X_f^T X_f)^{-1} X_f^T Y \\ = \; & c_I Q (Q^T X_0^T X_0 Q)^{-1} Q^T X_0 Y \\ = \; & c_I Q Q^{-1} (X_0^T X_0)^{-1} Q^{-T} Q^T X_0^T Y \\ = \; & c_I (X_0^T X_0)^{-1} X_0^T Y \\ = \; & c_I \hat{\beta}^{(0)} \end{aligned}$$

And

$$\begin{aligned} & Var(c_f \hat{\beta}^{(f)}) \\ = \; & c_f (X_f^T X_f)^{-1} c_f^T \times (Y - X_f \hat{\beta}^{(f)})^T (Y - X_f \hat{\beta}^{(f)})/n \\ = \; & c_0 Q (Q^T X_0^T X_0 Q)^{-1} Q^T c_0^T \times (Y - X_0 \hat{\beta}^{(0)})^T (Y - X_0 \hat{\beta}^{(0)})/n \\ = \; & c_0 (X_0^T X_0)^{-1} c_0^T \times (Y - X_0 \hat{\beta}^{(0)})^T (Y - X_0 \hat{\beta}^{(0)})/n \\ = \; & Var(c_0 \hat{\beta}^{(0)}) \end{aligned}$$

A.3 Proof of theorems in chapter IV

Lemma A.6. *For any random variable X , x_1, x_2, let $x_{min} := \min\{x_1, x_2\}, x_{max} := \max\{x_1, x_2\}$, we have $P(X \geq x_1) \geq P(X \geq x_2) - P(x_{min} \leq X \leq x_{max})$.*

Proof.

$$\begin{aligned}
P(X \geq x_2) &= P(X \geq x_2, x_2 \geq x_1) + P(X \geq x_2, x_2 < x_1) \\
&\leq P(X \geq x_1, x_2 \geq x_1) + P(X \geq x_1, x_2 < x_1) + P(x_2 \leq X \leq x_1, x_2 < x_1) \\
&= P(X \geq x_1) + P(x_2 \leq X \leq x_1, x_2 < x_1) \\
&\leq P(X \geq x_1) + P(x_{min} \leq X \leq x_{max})
\end{aligned}$$

Thus we have $P(X \geq x_1) \geq P(X \geq x_2) - P(x_{min} \leq X \leq x_{max})$. □

Proof of Theorem IV.2: For notation convenience we define $K = \arg\max_i \theta_i$ (i.e. $\forall i,\ \theta_K \geq \theta_i$).

$$\begin{aligned}
& P(K \in \hat{\mathcal{B}}) \\
= \ & P(\frac{\hat{\theta}_K - \hat{\theta}_1}{\hat{\sigma}_{k1}} \geq -\hat{c}_K, \cdots, \frac{\hat{\theta}_K - \hat{\theta}_{K-1}}{\hat{\sigma}_{K(K-1)}} \geq -\hat{c}_K) \\
= \ & P(w_{K1} \geq -\hat{c}_K - \frac{\theta_k - \theta_1}{\hat{\sigma}_{K1}}, \cdots, w_{K(K-1)} \geq -\hat{c}_K - \frac{\theta_k - \theta_N}{\hat{\sigma}_{K(K-1)}}) \\
\geq \ & P(w_{K1} \geq -\hat{c}_K, \cdots, w_{K(K-1)} \geq -\hat{c}_K) \\
\geq \ & P(w_{K1} \geq -c_K, w_{K2} \geq -\hat{c}_K, \cdots, w_{K(K-1)} \geq -\hat{c}_K) \\
& -P(\min\{\hat{c}_K, c_K\} \leq w_{K1} \leq \max\{\hat{c}_K, c_K\}, w_{K2} \geq -\hat{c}_K, \cdots, w_{K(K-1)} \geq -\hat{c}_K) \\
\geq \ & P(w_{K1} \geq -c_K, w_{K2} \geq -\hat{c}_K, \cdots, w_{K(K-1)} \geq -\hat{c}_K) \\
& \qquad -P(\min\{\hat{c}_K, c_K\} \leq w_{K1} \leq \max\{\hat{c}_K, c_K\}) \\
\cdots & \\
\geq \ & P(w_{K1} \geq -c_K, \cdots, w_{K(K-1)} \geq -c_K) - P(\min\{\hat{c}_K, c_K\} \leq w_{K1} \leq \max\{\hat{c}_K, c_K\}) \\
& \qquad - \cdots - P(\min\{\hat{c}_K, c_K\} \leq w_{K(K_1)} \leq \max\{\hat{c}_K, c_K\}) \\
= \ & 1 - \alpha + o_p(1).
\end{aligned}$$

The last equality follows from $P(w_{K1} \geq -c_K, \cdots, w_{K(K-1)} \geq -c_K) = 1 - \alpha$ and $\hat{c}_K \to_p c_K$ (Lemma A.6).

Proof of Theorem IV.3: We denote the pdf of Y_i as $f_i(y_i)$ for $i = 1, \cdots, N$. Let $\Phi(\cdot)$ denote the cdf of standard normal distribution.

When $\delta_{Ki} = -\delta_{iK}$ is large, we have

$$\begin{aligned}
& \lim_{\delta_{Ki}\to\infty} \frac{P(W_{i1} \le Y_1, W_{i2} \le Y_2, \cdots, W_{iK} \le Y_K - \frac{\delta_{Ki}}{\sigma_{Ki}})}{P(W_{i1} \le Y_1, W_{i2} \le Y_2, \cdots, W_{iK} \le Y_K)\exp(-\delta_{Ki})} \\
< & \lim_{\delta_{Ki}\to\infty} \frac{P(W_{iK} \le Y_K - \frac{\delta_{Ki}}{\sigma_{Ki}})}{P(W_{i1} \le Y_1, W_{i2} \le Y_2, \cdots, W_{iK} \le Y_K)\exp(-\delta_{Ki})} \\
= & \frac{1}{P(W_{i1} \le Y_1, W_{i2} \le Y_2, \cdots, W_{iK} \le Y_K)} \lim_{\delta_{Ki}\to\infty} \int_0^\infty \frac{\Phi(y_K - \delta_{Ki}/\sigma_{Ki})}{\exp(-\delta_{Ki})} f_K(y_K)\mathrm{d}y_K \\
= & \frac{1}{P(W_{i1} \le Y_1, W_{i2} \le Y_2, \cdots, W_{iK} \le Y_K)} \int_0^\infty \lim_{\delta_{Ki}\to\infty} \frac{\Phi(y_K - \delta_{Ki}/\sigma_{Ki})}{\exp(-\delta_{Ki})} f_K(y_K)\mathrm{d}y_K \\
= & 0.
\end{aligned}$$

From the third line to the fourth line we can validly change the order of integration and limit because $\frac{\Phi(y_K - \delta_{Ki}/\sigma_{Ki})}{\exp(-\delta_{Ki})} f_K(y_K)$ is bounded and thus absolutely continuous. Thus, there exists $M > 0$ such that

$$\begin{aligned}
& P(W_{i1} \le Y_1, W_{i2} \le Y_2, \cdots, W_{iK} \le Y_K - \frac{\delta_{Ki}}{\sigma_{Ki}}) \\
& \qquad \le P(W_{i1} \le Y_1, W_{i2} \le Y_2, \cdots, W_{iK} \le Y_K)\exp(-\delta_{Ki}) \quad \forall \delta_{Ki} \ge M.
\end{aligned} \tag{A.2}$$

Define $g(\delta_{Ki}) := \frac{\mathrm{d}}{\mathrm{d}\delta_{Ki}} P(W_{i1} \le Y_1, W_{i2} \le Y_2, \cdots, W_{iK} \le Y_K - \frac{\delta_{Ki}}{\sigma_{Ki}}) < 0$. Then, normality of W_i implies that $g(\delta_{Ki})$ is a smooth function of δ_{Ki} and $g(\delta_{Ki}) < 0$, for all $\delta_{Ki} \ge 0$. Thus, there exists a constant $\zeta_0 > 0$ independent of δ_{Ki}, such that $g(\delta_{Ki}) \le -\zeta_0$ for all $\delta_{Ki} \in [0, M]$.

Table A.1: Simulation SMART design Example 1: Inference about the parameters β using IPW, AIPW and AIPW$_m$ where the latter represents the misspecified scenario.

Parameter	n=100						n=400					
	IPW		AIPW		AIPW$_m$		IPW		AIPW		AIPW$_m$	
	Bias	S.D.	Bias	S.D	Bias	S.D	Bias	S.D.	Bias	S.D	Bias	S.D
β_0	0.010	0.24	0.002	0.23	0.007	0.24	0.004	0.12	0.000	0.12	0.007	0.12
β_1	0.001	0.24	0.002	0.18	0.005	0.18	0.011	0.12	0.002	0.09	0.002	0.10
β_2	0.002	0.17	0.003	0.07	0.002	0.10	0.000	0.08	0.002	0.04	0.004	0.05
θ_1	0.013	0.41	0.007	0.32	0.014	0.39	0.015	0.21	0.004	0.16	0.013	0.20
θ_2	0.011	0.33	0.003	0.27	0.004	0.31	0.007	0.17	0.000	0.14	0.009	0.15
θ_3	0.009	0.41	0.003	0.32	0.010	0.39	0.015	0.21	0.000	0.16	0.005	0.20
θ_4	0.007	0.33	0.003	0.27	0.000	0.31	0.007	0.17	0.004	0.14	0.001	0.15

For $\delta_{Ki} \in [0, M]$, we have

$$\begin{aligned}
&\frac{\log \frac{P(W_{i1}\le Y_1, W_{i2}\le Y_2, \cdots, W_{iK}\le Y_K)}{P(W_{i1}\le Y_1, W_{i2}\le Y_2, \cdots, W_{iK}\le Y_K - \frac{\delta_{Ki}}{\sigma_{Ki}})}}{\delta_{Ki}} \\
= \;& \frac{\mathrm{d}}{\mathrm{d}\delta_{Ki}} \log P(W_{i1} \le Y_1, W_{i2} \le Y_2, \cdots, W_{iK} \le Y_K - \frac{\delta_{Ki}}{\sigma_{Ki}}) \Big|_{\delta_{Ki}=\eta\in[0,\delta_{Ki}]\subset[0,M]} \\
= \;& -g(\eta)/P(W_{i1} \le Y_1, W_{i2} \le Y_2, \cdots, W_{iK} \le Y_K) \\
> \;& -g(\eta) > \zeta_0.
\end{aligned}$$

Thus,

$$\begin{aligned}
&P(W_{i1} \le Y_1, W_{i2} \le Y_2, \cdots, W_{iK} \le Y_K - \frac{\delta_{Ki}}{\sigma_{Ki}}) \\
&\qquad \le P(W_{i1} \le Y_1, W_{i2} \le Y_2, \cdots, W_{iK} \le Y_K) \exp(-\zeta\delta_{Ki}), \quad \forall \delta_{Ki} \in [0, M],
\end{aligned} \tag{A.3}$$

where $\zeta = \min\{1, \zeta_0\}$. Combining (A.2) and (A.3) completes the proof.

A.4 Tables for chapter IV

Here, we present the supplementary Tables from Sections 4.5 and 4.6

Table A.2: Simulation SMART design Example 2: Inference about the parameters β using IPW, AIPW and AIPW$_m$ where the latter represents the misspecified scenario.

Parameter	n=100						n=400					
	IPW		AIPW		AIPW$_m$		IPW		AIPW		AIPW$_m$	
	Bias	S.D.	Bias	S.D	Bias	S.D	Bias	S.D.	Bias	S.D	Bias	S.D
β_0	0.004	0.35	0.001	0.30	0.005	0.30	0.003	0.18	0.001	0.15	0.002	0.15
β_1	0.004	0.75	0.003	0.40	0.002	0.48	0.004	0.37	0.001	0.20	0.002	0.23
β_2	0.010	0.81	0.005	0.30	0.003	0.47	0.000	0.40	0.002	0.14	0.004	0.23
β_3	0.011	0.81	0.008	0.30	0.008	0.49	0.010	0.40	0.005	0.14	0.001	0.23
β_4	0.013	0.81	0.006	0.31	0.001	0.47	0.005	0.40	0.006	0.14	0.003	0.24
θ_1	0.004	0.35	0.001	0.30	0.005	0.30	0.003	0.18	0.001	0.15	0.002	0.15
θ_2	0.008	0.66	0.004	0.37	0.007	0.46	0.007	0.32	0.002	0.18	0.004	0.21
θ_3	0.018	0.66	0.009	0.38	0.010	0.45	0.007	0.33	0.004	0.19	0.008	0.21
θ_4	0.019	0.66	0.012	0.38	0.015	0.48	0.017	0.33	0.007	0.19	0.005	0.22
θ_5	0.021	0.66	0.010	0.38	0.006	0.45	0.012	0.33	0.008	0.18	0.007	0.21

A.5 Discussion and tables for the simulation results in section 4.8

In this section we present the description and results in each of the four scenarios in section 4.8. They are summarized by Rong Zhou.

For all the tables, we denote S_{ACI} as the expected set size for the ACI method, S_{MCB} as the expected set size for the MCB method, P_{ACI} as the probability of containing the true best DTR for the ACI method, and P_{MCB} as the probability of containing the true best DTR for the MCB method

A.5.0.1 Scenario One

In scenario one, all eight DTRs are equally optimal. In this scenario, different sample sizes in each setting are used. In addition, slightly altered γ_3 values are assigned as well, and as γ_3 changes from 0 to positive values, the optimal DTRs changes to when the first stage treatment selection is $A_1 = 1$. In Table 1, We find that the MCB method performs better than the modified ACI method.

Table A.3: Extend trial: Inference about the parameters β using IPW and AIPW.

Parameter	IPW		AIPW	
	Est.	S.D.	Est.	S.D
β_0	8.86	0.45	8.84	0.47
β_1	-0.99	0.45	-0.90	0.44
β_2	-0.24	0.34	-0.09	0.27
β_3	-0.07	0.28	-0.21	0.13
θ_1	7.56	0.76	7.65	0.67
θ_2	7.71	0.74	8.06	0.67
θ_3	8.05	0.71	7.83	0.70
θ_4	8.19	0.69	8.24	0.70
θ_5	9.53	0.81	9.44	0.76
θ_6	9.68	0.80	9.85	0.77
θ_7	10.02	0.83	9.62	0.70
θ_8	10.17	0.82	10.03	0.72

Table A.4: Results of Scenario One

Number of iteration, Sample size	γ vector setting	Best DTR index	P_{ACI}	P_{MCB}	S_{ACI}	S_{MCB}
10000 150	(0,0,0,0,0,0,0,0)	1-8	0.96 0.97 0.96 0.96 0.96 0.96 0.96 0.96	0.96 0.94 0.94 0.94 0.94 0.98 0.97 0.94	7.70	7.52
10000 300	(0,0,0,0,0,0,0,0)	1-8	0.96 0.96 0.96 0.97 0.97 0.96 0.96 0.96	0.94 0.95 0.95 0.95 0.94 0.94 0.94 0.95	7.70	7.56
10000 600	(0,0,0,0,0,0,0,0)	1-8	0.96 0.96 0.96 0.97 0.97 0.96 0.96 0.96	0.94 0.95 0.95 0.95 0.95 0.95 0.95 0.95	7.71	7.57
2000 150	(0,0,0.01,0,0,0,0)	1-4	0.97 0.97 0.96 0.96	0.94 0.95 0.95 0.94	7.70	7.47
2000 150	(0,0,0.05,0,0,0,0)	1-4	0.97 0.97 0.96 0.96	0.95 0.96 0.96 0.96	7.69	7.37
2000 150	(0,0,0.1,0,0,0,0)	1-4	0.97 0.97 0.96 0.96	0.96 0.97 0.97 0.97	7.65	7.03
2000 150	(0,0,0.2,0,0,0,0)	1-4	0.97 0.97 0.96 0.96	0.97 0.97 0.97 0.97	7.33	5.76

2000 150	(0,0,0.3,0,0,0,0)	1-4	0.97 0.97 0.96 0.96	0.97 0.97 0.97 0.97	6.47	4.52

A.5.0.2 Scenario Two

In scenario two, the final outcomes are different by the second stage decision A_2 if the first stage decision A_1 is 1, but are the same if the first stage decision A_1 is -1. In Table 2, we find that both methods have similar performance, with a slightly better result from MCB.

Table A.5: Results of Scenario Two

Iteration, Sample size	γ vector setting	Best DTR index	P_{ACI}	P_{MCB}	S_{ACI}	S_{MCB}
10000 150	(0,0,-0.5,0,0.5,0,0.5)	1 5 6 7 8	0.98 0.96 0.96 0.96 0.96	0.97 0.96 0.96 0.96 0.96	4.83	4.82
10000 300	(0,0,-0.5,0,0.5,0,0.5)	1 5 6 7 8	0.98 0.96 0.96 0.96 0.96	0.97 0.96 0.97 0.97 0.97	4.83	4.83
10000 600	(0,0,-0.5,0,0.5,0,0.5)	1 5 6 7 8	0.99 0.97 0.96 0.96 0.96	0.97 0.97 0.97 0.97 0.97	4.84	4.83
2000 150	(0,0,-0.49,0,0.51,0,0.51)	1	0.99	0.98	4.83	4.80
2000 150	(0,0,-0.5,0,0.5,0,0.5)	1 4	0.99 0.97	0.97 0.97	4.82	4.81

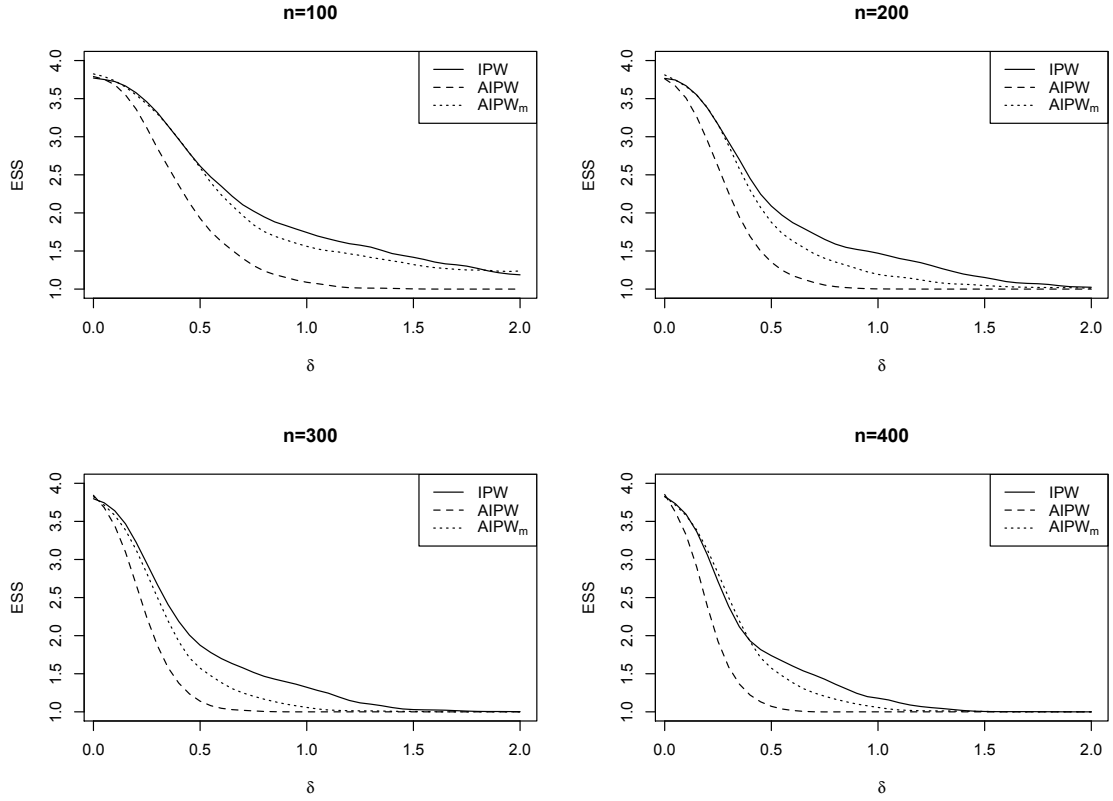

Figure A.1: Simulation SMART design Example 1: The vertical axes is the estimated set (of best) size (ESS) and horizontal axes is the difference between the best and the second best EDTR.

A.5.0.3 Scenario Three

In scenario three, responders to the first stage have the same final outcomes to A_2, but non-responders have different expected final outcomes to A_2. In Table 3, we find that the MCB method performs better than the modified ACI method.

Table A.6: Results of Scenario Three

Iteration, Sample size	γ vector setting	Best DTR index	P_{ACI}	P_{MCB}	S_{ACI}	S_{MCB}

2000 150	(0,0,0,0,0.5,-0.5,0)	1 3 5 7	0.98 0.97 0.97 0.97	0.97 0.97 0.96 0.96	3.89	3.88
2000 150	(0,0,0.01,0,0.5,-0.5,0)	1 3	0.98 0.97	0.98 0.97	3.89	3.88
2000 150	(0,0,0.05,0,0.5,-0.5,0)	1 3	0.99 0.98	0.98 0.98	3.88	3.85
2000 150	(0,0,0.1,0,0.5,-0.5,0)	1 3	0.99 0.98	0.99 0.98	3.85	3.77
2000 150	(0,0,0.2,0,0.5,-0.5,0)	1 3	0.99 0.98	0.99 0.99	3.60	3.35
2000 150	(0,0,0.3,0,0.5,-0.5,0)	1 3	0.99 0.98	0.99 0.99	3.16	2.79

A.5.0.4 Scenario Four

In scenario four, we manipulated the γ vector such that only one or two optimal DTRs will be obtained in this case, and there is no non-regularity. In Table 4, we find that the modified ACI method performs better than the MCB method. Because there is no non-regularity in this scenario, the modified ACI method will not be conservative compared to its performance in the first three scenarios. Also, since there are only two optimal DTRs in this scenario, and the MCB method still uses the critical value for comparing all eight DTRs when it chooses a set in this case, the result of the MCB method will be conservative.

Table A.7: Results of Scenario Four

Iteration, Sample size	γ vector setting	Best DTR index	P_{ACI}	P_{MCB}	S_{ACI}	S_{MCB}
2000 150	(0,0,-1,0,-1,0,-0.5)	4 8	0.98 0.98	0.99 0.99	1.96	1.98
2000 150	(0,0,-0.99,0,-1,0,-0.5)	4	0.98	0.99	1.96	1.98

2000 150	(0,0,-0.45,0,-1,0,-0.5)	4	0.99	0.99	1.95	1.98
2000 150	(0,0,-0.4,0,-1,0,-0.5)	4	1.00	1.00	1.93	1.96
2000 150	(0,0,-0.51,0,-1,0,-0.5)	8	0.98	0.99	1.96	1.98
2000 150	(0,0,-0.55,0,-1,0,-0.5)	8	0.99	1.00	1.95	1.98
2000 150	(0,0,-0.6,0,-1,0,-0.5)	8	1.00	1.00	1.92	1.96

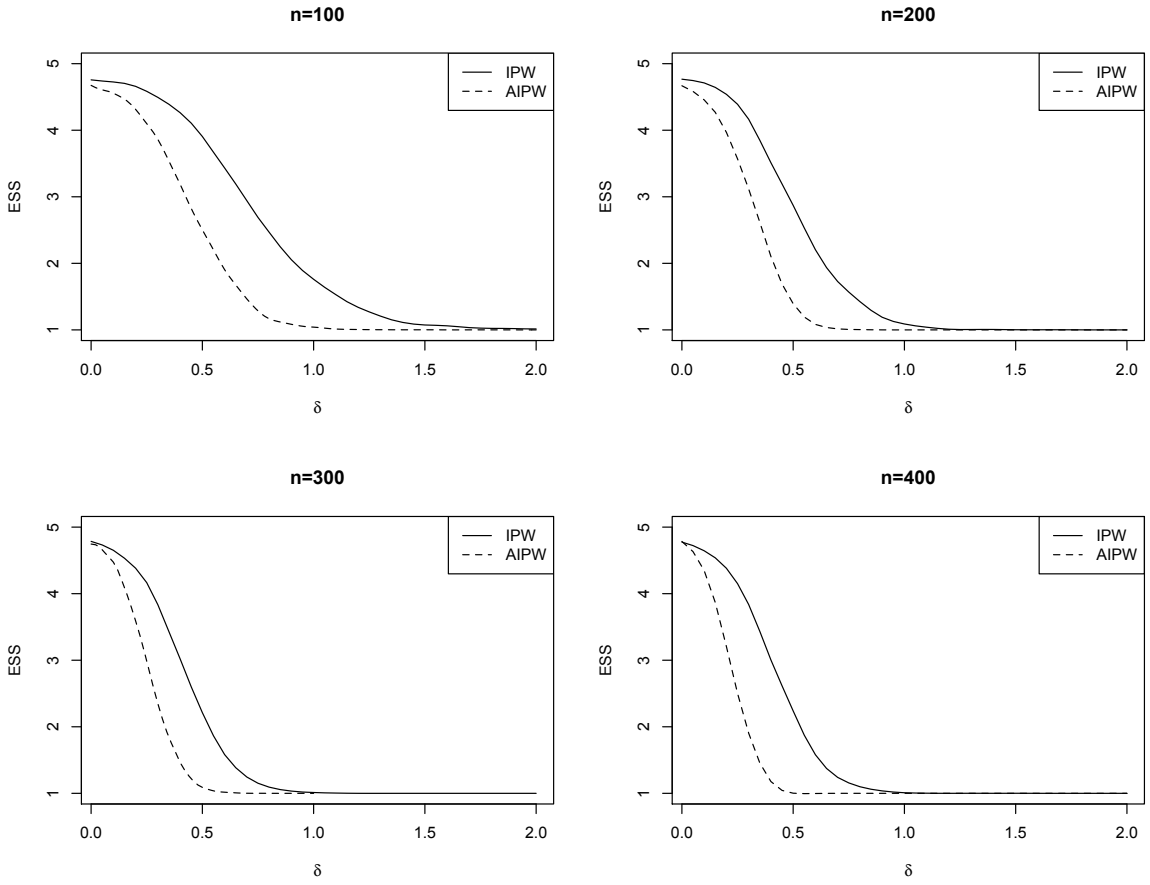

Figure A.2: Simulation SMART design Example 2: The vertical axis are the estimated set (of best) size (ESS) and horizontal axes are the difference between the best and the second best EDTR.

BIBLIOGRAPHY

BIBLIOGRAPHY

Aitchison, J., and I. Dunsmore (1968), Linear-loss interval estimation of location and scale parameters, *Biometrika*, *55*(1), 141–148.

Almirall, D., I. Nahum-Shani, N. E. Sherwood, and S. A. Murphy (2014), Introduction to smart designs for the development of adaptive interventions: with application to weight loss research, *Translational behavioral medicine*, *4*(3), 260–274.

Bang, H., and J. Robins (2005), Doubly robust estimation in missing data and causal inference models, *Biometrics*, *61*(4), 962–972.

Barnett, G. O., J. J. Cimino, J. A. Hupp, and E. P. Hoffer (1987), Dxplain: an evolving diagnostic decision-support system, *Jama*, *258*(1), 67–74.

Berger, R. L. (1989), Uniformly more powerful tests for hypotheses concerning linear inequalities and normal means, *Journal of the American Statistical Association*, *84*(405), 192–199.

Berger, R. L., et al. (1979), Minimax subset selection for loss measured by subset size, *The Annals of Statistics*, *7*(6), 1333–1338.

Bjørnstad, J. F. (1986), Asymptotic consistency for subset selection procedures satisfying the p-condition, *Journal of statistical planning and inference*, *13*, 319–335.

Blatt, D., S. Murphy, and J. Zhu (2004), A-learning for approximate planning, *Ann Arbor*, *1001*, 48,109–2122.

Blois, M., M. Tuttle, and D. Sherertz (1981), Reconsider: A program for generating differential diagnoses*., in *Proceedings/the... Annual Symposium on Computer Application [sic] in Medical Care. Symposium on Computer Applications in Medical Care*, pp. 263–268, American Medical Informatics Association.

Bonferroni, C. E. (1936), *Teoria statistica delle classi e calcolo delle probabilita*, Libreria internazionale Seeber.

Box, G. E., et al. (1954a), Some theorems on quadratic forms applied in the study of analysis of variance problems, i. effect of inequality of variance in the one-way classification, *The annals of mathematical statistics*, *25*(2), 290–302.

Box, G. E., et al. (1954b), Some theorems on quadratic forms applied in the study of analysis of variance problems, ii. effects of inequality of variance and of correlation between errors in the two-way classification, *The Annals of Mathematical Statistics*, *25*(3), 484–498.

Casella, G., and J. T. Hwang (1991), Evaluating confidence sets using loss functions, *Statistica Sinica*, pp. 159–173.

Chakraborty, B., and E. Moodie (2013), *Statistical Methods for Dynamic Treatment Regimes*, Springer.

Chakraborty, B., and S. A. Murphy (2014), Dynamic treatment regimes, *Annual Review of Statistics and its Application*, *1*, 447–464.

Chakraborty, B., S. Murphy, and V. Strecher (2009), Inference for non-regular parameters in optimal dynamic treatment regimes, *Statistical methods in medical research*.

Chakraborty, B., E. B. Laber, and Y. Zhao (2013), Inference for optimal dynamic treatment regimes using an adaptive m-out-of-n bootstrap scheme, *Biometrics*, *69*(3), 714–723.

Cohen, A., W. E. Strawderman, et al. (1973), Admissibility implications for different criteria in confidence estimation, *The Annals of Statistics*, *1*(2), 363–366.

Davidian, M., A. Tsiatis, and S. Leon (2005), Semiparametric estimation of treatment effect in a pretest–posttest study with missing data, *Statistical Science*, *20*(3), 261–301.

Dunn, O. J. (1959), Estimation of the medians for dependent variables, *The Annals of Mathematical Statistics*, pp. 192–197.

Dunn, O. J. (1961), Multiple comparisons among means, *Journal of the American Statistical Association*, *56*(293), 52–64.

Dunnett, C. W. (1955), A multiple comparison procedure for comparing several treatments with a control, *Journal of the American Statistical Association*, *50*(272), 1096–1121.

Edwards, D. G., and J. C. Hsu (1983), Multiple comparisons with the best treatment, *Journal of the American Statistical Association*, *78*(384), 965–971.

Fard, M. M., and J. Pineau (2009), Mdps with non-deterministic policies, in *Advances in neural information processing systems*, pp. 1065–1072.

Fortney, J. C., et al. (2010), A web-based clinical decision support system for depression care management, *The American journal of managed care*, *16*(11), 849.

Gupta, S. S. (1956), On a decision rule for a problem in ranking means, Ph.D. thesis, University of North Carolina at Chapel Hill.

Gupta, S. S. (1965), On some multiple decision (selection and ranking) rules, *Technometrics*, *7*(2), 225–245.

Gupta, S. S., and D.-Y. Huang (1980), An essentially complete class of multiple decision procedures, *Journal of Statistical Planning and Inference*, *4*(2), 115–121.

Gupta, S. S., and H.-M. Yang (1985), Bayes-p subset selection procedures for the best population, *Journal of statistical planning and inference*, *12*, 213–233.

Hansen, P. R., A. Lunde, and J. M. Nason (2011), The model confidence set, *Econometrica*, *79*(2), 453–497.

Henderson, R., P. Ansell, and D. Alshibani (2010), Regret-regression for optimal dynamic treatment regimes, *Biometrics*, *66*(4), 1192–1201.

Hernán, M. Á., B. Brumback, and J. M. Robins (2000), Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men, *Epidemiology*, *11*(5), 561–570.

Hirano, K., G. W. Imbens, and G. Ridder (2003), Efficient estimation of average treatment effects using the estimated propensity score, *Econometrica*, *71*(4), 1161–1189.

Hooper, P. M. (1982), Invariant confidence sets with smallest expected measure, *The Annals of Statistics*, pp. 1283–1294.

Horrace, W. C., P. Schmidt, et al. (2000), Multiple comparisons with the best, with economic applications, *Journal of Applied Econometrics*, *15*(1), 1–26.

Hsu, J. (1996), *Multiple comparisons: theory and methods*, CRC Press.

Hsu, J. C., et al. (1981), Simultaneous confidence intervals for all distances from the "Best", *The Annals of Statistics*, *9*(5), 1026–1034.

Hsu, J. C., et al. (1984), Constrained simultaneous confidence intervals for multiple comparisons with the best, *The Annals of Statistics*, *12*(3), 1136–1144.

Hsu, T.-A. (1985), Some optimal subset selection procedures with applications, *Sankhyā: The Indian Journal of Statistics, Series B*, pp. 206–215.

Hwang, J. T., and G. Casella (1982), Minimax confidence sets for the mean of a multivariate normal distribution, *The Annals of Statistics*, pp. 868–881.

Joshi, V. (1969), Admissibility of the usual confidence sets for the mean of a univariate or bivariate normal population, *The Annals of Mathematical Statistics*, *40*(3), 1042–1067.

Keselman, H., R. Cribbie, and B. Holland (1999), The pairwise multiple comparison multiplicity problem: An alternative approach to familywise and comparison wise Type I error control., *Psychological Methods*, *4*(1), 58.

Laber, E. B., M. Qian, D. J. Lizotte, W. E. Pelham, and S. A. Murphy (2010), Statistical inference in dynamic treatment regimes, *arXiv preprint arXiv:1006.5831*.

Laber, E. B., D. J. Lizotte, and B. Ferguson (2014a), Set-valued dynamic treatment regimes for competing outcomes, *Biometrics*, *70*(1), 53–61.

Laber, E. B., D. J. Lizotte, M. Qian, W. E. Pelham, and S. A. Murphy (2014b), Dynamic treatment regimes: Technical challenges and applications, *Electronic journal of statistics*, *8*(1), 1225.

Lavori, P. W., and R. Dawson (2000), A design for testing clinical strategies: biased adaptive within-subject randomization, *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *163*(1), 29–38.

Lavori, P. W., and R. Dawson (2008), Adaptive treatment strategies in chronic disease, *Annual review of medicine*, *59*, 443.

Lavori, P. W., R. Dawson, and A. J. Rush (2000), Flexible treatment strategies in chronic disease: clinical and research implications, *Biological Psychiatry*, *48*(6), 605–614.

Lehmann, E. (1952), Testing multiparameter hypotheses, *The Annals of Mathematical Statistics*, pp. 541–552.

Lei, H., I. Nahum-Shani, K. Lynch, D. Oslin, and S. Murphy (2012), A smart design for building individualized treatment sequences, *Annual review of clinical psychology*, *8*.

Lunceford, J. K., M. Davidian, and A. A. Tsiatis (2002), Estimation of survival distributions of treatment policies in two-stage randomization designs in clinical trials, *Biometrics*, *58*(1), 48–57.

Miller, R., and F. Masarie Jr (1989), Use of the quick medical reference (qmr) program as a tool for medical education., *Methods of information in medicine*, *28*(4), 340–345.

Moodie, E. E., and T. S. Richardson (2010), Estimating optimal dynamic regimes: Correcting bias under the null, *Scandinavian Journal of Statistics*, *37*(1), 126–146.

Moodie, E. E., T. S. Richardson, and D. A. Stephens (2007), Demystifying optimal dynamic treatment regimes, *Biometrics*, *63*(2), 447–455.

Murphy, S., M. Van Der Laan, and J. Robins (2001), Marginal mean models for dynamic regimes, *Journal of the American Statistical Association*, *96*(456), 1410–1423.

Murphy, S. A. (2003), Optimal dynamic treatment regimes, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *65*(2), 331–355.

Murphy, S. A. (2005a), An experimental design for the development of adaptive treatment strategies, *Statistics in medicine*, *24*(10), 1455–1481.

Murphy, S. A. (2005b), A generalization error for q-learning, *Journal of machine learning research: JMLR*, *6*, 1073.

Nahum-Shani, I., M. Qian, D. Almirall, W. E. Pelham, B. Gnagy, G. A. Fabiano, J. G. Waxmonsky, J. Yu, and S. A. Murphy (2012a), Experimental design and primary data analysis methods for comparing adaptive interventions., *Psychological methods*, *17*(4), 457–477.

Nahum-Shani, I., M. Qian, D. Almirall, W. E. Pelham, B. Gnagy, G. A. Fabiano, J. G. Waxmonsky, J. Yu, and S. A. Murphy (2012b), Q-learning: A data analysis method for constructing adaptive interventions., *Psychological methods*, *17*(4), 478.

Orellana, L., A. Rotnitzky, and J. M. Robins (2010), Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, part i: main content, *The International Journal of Biostatistics*, *6*(2).

Perlman, M. D., L. Wu, et al. (1999), The emperors new tests, *Statistical Science*, *14*(4), 355–369.

Robins, J. (1986), A new approach to causal inference in mortality studies with a sustained exposure periodapplication to control of the healthy worker survivor effect, *Mathematical Modelling*, *7*(9), 1393–1512.

Robins, J., L. Orellana, and A. Rotnitzky (2008), Estimation and extrapolation of optimal treatment and testing strategies, *Statistics in medicine*, *27*(23), 4678–4721.

Robins, J., A. Rotnitzky, et al. (2014), Discussion of dynamic treatment regimes: Technical challenges and applications, *Electronic Journal of Statistics*, *8*, 1273–1289.

Robins, J. M. (1989), The analysis of randomized and non-randomized aids treatment trials using a new approach to causal inference in longitudinal studies, *Health service research methodology: a focus on AIDS*, *113*, 159.

Robins, J. M. (1993), Information recovery and bias adjustment in proportional hazards regression analysis of randomized trials using surrogate markers, in *Proceedings of the Biopharmaceutical Section, American Statistical Association*, vol. 24, p. 3, American Statistical Association.

Robins, J. M. (1997), Causal inference from complex longitudinal data, in *Latent variable modeling and applications to causality*, pp. 69–117, Springer.

Robins, J. M. (1999), Association, causation, and marginal structural models, *Synthese*, *121*(1), 151–179.

Robins, J. M. (2004), Optimal structural nested models for optimal sequential decisions, in *Proceedings of the Second Seattle Symposium in Biostatistics*, pp. 189–326, Springer.

Robins, J. M., A. Rotnitzky, and L. P. Zhao (1995), Analysis of semiparametric regression models for repeated outcomes in the presence of missing data, *Journal of the American Statistical Association*, *90*(429), 106–121.

Robins, J. M., M. A. Hernan, and B. Brumback (2000), Marginal structural models and causal inference in epidemiology, *Epidemiology*, *11*(5), 550–560.

Roth, A. J. (1978), A new procedure for selecting a subset containing the best normal population, *Journal of the American Statistical Association*, *73*(363), 613–617.

Roy, S. N. (1953), On a heuristic method of test construction and its use in multivariate analysis, *The Annals of Mathematical Statistics*, pp. 220–238.

Saville, D. J. (1990), Multiple comparison procedures: the practical solution, *The American Statistician*, *44*(2), 174–180.

Scheffe, H. (1953), A method for judging all contrasts in the analysis of variance, *Biometrika*, *40*(1-2), 87–110.

Seal, K. (1955), On a class of decision procedures for ranking means of normal populations, *The Annals of Mathematical Statistics*, pp. 387–398.

Seidler, J., J. Vondráček, and I. Saxl (2000), The life and work of zbyněk šidák (1933–1999), *Applications of Mathematics*, *45*(5), 321–336.

Sen, P. K. (2007), Union–intersection principle and constrained statistical inference, *Journal of Statistical Planning and Inference*, *137*(11), 3741–3752.

Shaffer, J. P. (1995), Multiple hypothesis testing, *Annual review of psychology*, *46*, 561.

Šidák, Z. (1967), Rectangular confidence regions for the means of multivariate normal distributions, *Journal of the American Statistical Association*, *62*(318), 626–633.

Thall, P. F., R. E. Millikan, H.-G. Sung, et al. (2000), Evaluating multiple treatment courses in clinical trials, *Statistics in medicine*, *19*(8), 1011–1028.

Thall, P. F., H.-G. Sung, and E. H. Estey (2011), Selecting therapeutic strategies based on efficacy and death in multicourse clinical trials, *Journal of the American Statistical Association*.

Trivedi, M. H., J. K. Kern, B. D. Grannemann, K. Z. Altshuler, and P. Sunderajan (2014), A computerized clinical decision support system as a means of implementing depression guidelines.

Tukey, J. W. (1953), *The Problem of Multiple Comparisons: Introduction and Parts A, B, and C*, B [Princeton University].

van der Laan, M., and J. Robins (2003), *Unified methods for censored longitudinal data and causality*, Springer Verlag.

Wahed, A. S., and A. A. Tsiatis (2004), Optimal estimator for the survival distribution and related quantities for treatment policies in two-stage randomization designs in clinical trials, *Biometrics*, *60*(1), 124–133.

Wahed, A. S., and A. A. Tsiatis (2006), Semiparametric efficient estimation of survival distributions in two-stage randomisation designs in clinical trials with censored data, *Biometrika*, *93*(1), 163–177.

Warner Jr, H. (1989), Iliad: moving medical decision-making into new frontiers., *Methods of information in medicine*, *28*(4), 370–372.

Watkins, C. J., and P. Dayan (1992), Q-learning, *Machine learning*, *8*(3-4), 279–292.

Winkler, R. L. (1972), A decision-theoretic approach to interval estimation, *Journal of the American Statistical Association*, *67*(337), 187–191.

Zhao, Y., M. R. Kosorok, and D. Zeng (2009), Reinforcement learning design for cancer clinical trials, *Statistics in medicine*, *28*(26), 3294–3315.